BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

5-11-04

1634

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of: Xiao
Serial No.: 09/845,416 Group No.: 1634
Filed: 04/30/01 Examiner: B. Whiteman
Entitled: **DNA Sequences Encoding Dystrophin Minigenes And Methods Of Use Thereof**

PROTEST TRANSMITTAL LETTER

Assistant Commissioner for Patents
ATTN: Technology Center 1600 - Driector Doll
P.O. Box 1450
Alexandria, VA 22313-1450

CERTIFICATE OF MAILING UNDER 37 C.F.R. § 1.10

I hereby certify that this correspondence (along with any referred to as being attached or enclosed) are being deposisted with the U.S. Postal Service on **May 10, 2004** in an envelope as "EXPRESS MAIL POST OFFICE TO ADDRESSEE" service under 37 C.F.R. § 1.10, Express Mail Label Number **EL 992 784 638 US** addressed to: Assistant Commissioner for Patents, ATTN: Technology Center 1600 - Director Doll, P.O. Box 1450, Alexandria, VA 22313-1450

By: _____
Jennifer B. Xistris

REMARKS

This is a Protest for application no. 09/845,416, filed 04/30/01. Third Party Protestors believe no fee is required but if the Commissioner deems otherwise he is authorized to charge Deposit Accout No. 08-1290.

A copy of this Protest is also being forwarded on this day to counsel of record: David A. Einhorn, Anderson, Kill, & Olick, P.C., 1251 Avenue of the America, New York, NY 10020 in an envelope as "Express Mail Post Office to Addressee" service under 37 C.F.R. 1.10, Express Mail Label No. EL 658 779 192 US.

DATE: May 10, 2004 By: _____

Jason R. Bond
Registration No. 45,439
MEDLEN & CARROLL, LLP
440 Science Drive, Suite 203
Madison, Wisconsin 53711
608-218-6900

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:	Xiao		
Serial No.:	09/845,416	Group No.:	1634
Filed:	04/30/01	Examiner:	B. Whiteman
Entitled:	**DNA Sequences Encoding Dystrophin Minigenes and Methods of Use Thereof**		

Protest Under 37 CFR 1.291, and Request to Withdraw from Issue Under 37 CFR 1.313, Based on Recently Published 102(e) Reference

Assistant Commissioner for Patents
Attn: Technology Center 1600 - Director Doll
P.O. Box 1450
Alexandria, VA 22313-1450

CERTIFICATE OF MAILING UNDER 37 C.F.R. § 1.10

I hereby certify that this correspondence (along with any referred to as being attached or enclosed) are being deposited with the U.S. Postal Service on **May 7, 2004** in an envelope as "EXPRESS MAIL POST OFFICE TO ADDRESSEE" service under 37 C.F.R. 1.10, Express Mail Label No. **EL 992784638 US** addressed to Assistant Commissioner for Patents, P.O.Box 1450, Alexandria, VA, ATTN: John Doll and Brian Whiteman.

Dated: May 10, 2004

By: Jennifer B. Xistris

Examiner Whiteman and Director Doll:

The following communication is presented to protest pending Application 09/845,416 (Xiao Application) under 37 CFR 1.291, and to request that this Application be withdrawn from issuance at the initiative of the Office under 37 CFR 1.313. This Application was allowed on 4-28-04 (currently located in the Certification Division). To the extent the claims in this Application encompass mini-dystrophin genes and peptides with **4 rod repeats** (as found in the claims of the corresponding Xiao Patent Pub), these claims should be withdrawn from issuance and rejected in view of recently published Patent Pub. US2003/0216332 to Chamberlain et al. ("Chamberlain Patent Pub") which anticipates such claims under 35 USC 102(e).

I. 4 Rod Repeats

This Protest and request for Patent Office initiated withdrawal from issuance of the allowed Xiao claims is relevant to claims encompassing mini-dystrophins with **4 rod repeats**. The Xiao Patent Pub (US2003/0171312) claims will be used in this communication as a guide to discuss the applicability of the Chamberlain Patent Pub as an anticipating 102(e) reference for claims than encompass 4 rod repeats. This discussion will find use even if the allowed Xiao claims differ from the Xiao Patent Pub claims.

Claim 1 of the Xiao Patent Pub includes 4 rod repeats. This claim is reproduced below:

1. An isolated nucleotide sequence encoding a dystrophin minigene comprising:
(a) a N-terminal domain;
(b) **four** to six **rod repeats**;
(c) an HI domain of a dystrophin gene and an H4 domain of the dystrophin gene; and
(d) a cysteine-rich domain,

wherein the N-terminal domain is selected from a group consisting of a N-terminal domain of the dystrophin gene, a modified N-terminal domain of the dystrophin gene, and a N-terminal domain of a utrophin gene; the rod repeats are selected from a group consisting of rod repeats in the dystrophin gene, rod repeats in the utrophin gene, and rod repeats in a spectrin gene; the cysteine-rich domain is a cysteine-rich domain of the dystrophin gene or a cysteine-rich domain of the utrophin gene.

II. Chamberlain Patent Pub Anticipates Xiao Claims Under 102(e)

The Chamberlain Patent Pub anticipates any allowed Xiao claims that encompass 4 rod repeats under 102(e) as the Chamberlain Patent Pub discloses 4 rod repeat mini-dystrophins in an Application filed prior to the filing date of the Xiao priority filing date. As shown in the time-line below, even though the Xiao Application claims priority to a Provisional Application that predates the Chamberlain Provisional, the Xiao Provisional *does not disclose 4 rod repeat mini-dystrophins*. Indeed, it explicitly excludes them. The Chamberlain Provisional, however, discloses 4 rod repeat mini-dystrophins - thereby predating the Xiao disclosure of 4 rod repeat mini-dystrophins by almost 7 months. Consequently, the Chamberlain Patent Pub is a 102(e) reference with regard to any Xiao claims that encompass 4 rod repeat mini-dystrophins.

A. Overview/Time-Line

Presented below is a brief overview of time-line of the Xiao and Chamberlain Provisional and Regular patent applications that shows the Chamberlain Provisional was the first to disclose 4 rod repeat mini-dystrophins.

1. **April 28, 2000:** Xiao Provisional Application No. 60/200,777 is filed (at Tab A), and discloses 5 and 6 rod repeat mini-dystrophins, but **does not disclose 4 rod repeat mini-dystrophins**. This Application explicitly excludes 4 rod repeats.

2. **October 6, 2000:** Chamberlain Provisional Application No. 60/238,848 is filed (Tab B), and **discloses 4 rod repeat mini-dystrophins**.

3. **April 30, 2001:** Xiao Regular Application No. 09/845,416 is filed (Xiao Patent Pub is at Tab C), and disclosed 4 rod repeat mini-dystrophins for the first time (nearly 7 months after Chamberlain Provisional filing date).

4. **October 4, 2001:** Chamberlain Regular Application No. 10/149,736 is filed (Chamberlain Patent Pub is at Tab D) claiming priority to October 6, 2000 Provisional Application that discloses 4 rod repeat mini-dystrophins. Current status: active prosecution of 4 rod repeat mini-dystrophin claims.

5. **November 20, 2003:** Chamberlain Patent Pub is published. PAIR indicates that the last substantive Office Action in the Xiao prosecution was mailed June 30, 2003. Since the Chamberlain Patent Pub was not published until November of 2003, it appears that Examiner was not aware of this reference when the Xiao claims were allowed.

B. Xiao Provisional Does Not Disclose 4 Rod Repeats Mini-Dystrophins

The Xiao Provisional Application, filed April 28, 2000, does not disclose 4 rod repeat mini-dystrophins (See Xiao Provisional at Tab A). Four rod repeat mini-dystrophins were first introduced by Xiao in the Xiao Regular Application filed April 30, 2001. As such, claims encompassing 4 rod repeats are not entitled to the Xiao Provisional filing date, but instead, only receive the filing date of the Xiao Regular Application.

The lack of teaching in this Provisional is highlighted by affirmative statements about the need to retain at least 5 rod repeats in the mini-dystrophins. For example, the Xiao Provisional states:

> "To ensure sufficient physical flexibility of the protein, all of our mini-dystrophins still retain **at least five rod repeats** (R1, R2, R22, R23 & R24) and 2 hinges (H1 and H4) in the central rod domain (Fig. 1)." (*emphasis added,* page 56).
>
> "However, the mini-dystrophin genes reported here accomodated **at least 5 rod repeats** (R1, R2, R22, R23 & R24) and two hinges (H1 and H4). Therefore we hypothesized that the length of the central rod domain is the most critical factor, based on the fact that a major role of dystrophin is to crosslink the myofiber cytoskelton and plasma membrane and stabilize the structure during muscle contractions. **If the dystrophin is too short** to span the sliding distance between the cytoskeleton and plasma membrane during muscle contraction, the crosslink will be disrupted and the muscle membrane will become unstable and prone to mechanical damages." (*emphasis added,* page 60).

The lack of teaching of 4 rod repeats in the Xiao Provisional is also seen by comparing the Xiao Provisional to the Xiao Regular Application. For example, it is instructive to compare Figure 1 in the Xiao Provisional (no 4 rod repeats are disclosed) to Figure 1 in the Xiao Regular Application (three 4 rod repeat constructs have been inserted into the middle of the figure). The description of Figure 1 in each application also highlights the change from at least 5 rod repeats in the Xiao Provisional (deleting 19 of the 24 rod repeats from natural dystrophin) to the 4 rod repeats in the Xiao Regular Application (deleting 20 of the rod repeats from natural dystrophin):

> Xiao Provisional page 56: "We have created by rational design several mini-genes, in each deleting up to 3/4 of the central rod domain (19 rods and 2 hinges) and nearly the entire distal C-terminal domain (exons 71-78) (Fig. 1)."
>
> Xiao Patent Pub Col. 5, par. 61: "We have created minigenes in which up to 75% of the central rod domain (20 of the 24 rods; 2 of the 4 hinges), as well as nearly all of the C-terminal domain (exons 71-78), are deleted (FIG. 1)."

In light of the lack of teaching of 4 rod repeats in the Xiao Provisional, it is clear that claims that encompass 4 rod repeats in the Xiao Regular Application are not entitled to the priority date of the Xiao Provisional Application.

C. Xiao Claims Are Anticipated by the Chamberlain Patent Pub Under 102(e)

As detailed above, Xiao claims that encompass 4 rod repeats are only entitled to the April 30, 2001 filing date of the Xiao Regular Application, not the April 28, 2000 filing date of the Xiao Provisional Application. The Chamberlain Provisional, however, fully discloses 4 rod repeat mini-dystrophins and was filed October 6, 2000. As such, the Chamberlain Patent Pub is a 102(e) reference that pre-dates the Xiao Application for 4 rod repeats by almost 7 months.

i. Chamberlain Provisional Fully Discloses 4 Rod Repeats

Unlike the Xiao Provisional, the Chamberlain Provisional fully discloses 4 rod repeats. Discussion of 4 rod repeats is found throughout the 142 page Chamberlain Provisional. For example, 4 "spectrin-like repeats" (rod repeats) are taught at the following pages: page 3, lines 13 and 28; page 4, lines 3 and 26; page 5, lines 2, and 4-7 (SEQ IDs 39-41); page 23, lines 3-4; Example 2 on pages 51-54 (detailing construction of ΔR4-R23, ΔR2-R21+H3, and ΔR2-R1; all of which have 4 spectrin-like repeats); Example 3 on pages 55-58 (detailing construction of ΔR4-

R23-71-78, which is a 4 spectrin like repeats construct with a C-terminal deletion); Example 5 on page 63-70 shows testing of various constructs including many 4 spectrin-like repeat constructs; Example 6 on pages 70-71 (describing how a 4 spectrin-like construct could be inserted into AAV and used to treat DMD or BMD); Figures 12-14 show the nucleic acid sequences for ΔR4-R23 (micro-dys 1; SEQ ID NO:39), ΔR2-R21 (micro-dys 2; SEQ ID NO:40), and ΔR2-R21-H3 (micro-dys 3; SEQ ID NO:41); and Figure 27, which shows four 4 spectrin-like repeat constructs (the last 4 constructs in Figure 27). In light of the extensive disclosure of 4 spectrin-like (rod) repeats in the Chamberlain application, including Examples testing these constructs *in vivo* on dystrophic mouse models, it is clear that the Chamberlain Patent Pub is fully entitled to the priority date of the Chamberlain provisional for 4 rod repeat mini-dystropins.

ii. Chamberlain Patent Pub Anticipates Exemplary Claim 1

The Chamberlain Pat Pub, based on the disclosure of the Chamberlain Provisional, fully anticipates Claim 1 of the Xiao Pat Pub. Claim 1 of the Xiao Patent Pub recites a nucleic acid sequence encoding a mini-dystropin peptide with the following four elements:

(a) a N-terminal domain;
(b) **four** to six **rod repeats**;
(c) an HI domain of a dystrophin gene and an H4 domain of the dystrophin gene; and
(d) a cysteine-rich domain.

The Chamberlain Provisional teaches nucleic acid sequences encoding mini-dystropins with the four recited elements.

As described above, the Chamberlain Provisional describes constructs with 4 rod repeats. For example, the last four constructs in Figure 27 shows such 4 rod repeats. These constructs are labeled: R4-R23 (micro-dys 1), R2-R21 (micro-dys 2), R2-R21-H3 (micro-dys 3), and R4-R23-71-78. Each of these four constructs contain an "H1 domain" and an "H4 domain" as recited in Claim 1. These four constructs also contain a "CR" (cystein-rich domain), as well as an N-terminal domain (labeled ABD1 for actin-binding domain, which is the N-terminal domain; see, e.g, SEQ ID NO:6).

The four elements in Claim 1 are also taught in additional locations throughout the Chamberlain Provisional (i.e. not just in Figure 27). For example, the Chamberlain specification teaches the use of an "actin-binding domain" (aka N-terminal domain) throughout the summary of the invention and gives SEQ ID NO:6 as an example of a 756 nucleic acid sequence that

encodes an actin-binding domain (N-terminal domain). Figures 12-14 depicting SEQ ID NOs:39-41 also teach 4 rod repeat constructs with N-terminal domains (actin-binding domains).

The Chamberlain Provisional also teaches the use of H1 and H4 in mini-dystrophins. For example, page 6 lines 15-26; page 9, lines 18-29; and Figures 12-14 depicting SEQ ID NOs:39-41 (that teach 4 rod repeat constructs containing hinges 1 and 4). Cysteine-rich domains are also taught on page 5, lines 23-27; in SEQ ID NO:35, and in Figures 12-14 (teaching the nucleic acid sequence of 4 rod repeat constructs with cystein-rich domains).

The last part of Claim 1 specifies that the various elements (e.g. N-terminals, rod repeats, and cystein-rich domains) can come from various sources such as dystrophin, modified dystrophin, utrophin, and spectrin. The Chamberlain Provisional not only teaches dystrophin, but also contains a section entitled "Variants and Homologs of Dystrophin" which teaches, for example, variants of dystrophin, and dystrophin homologs such as utrophin and alpha-actinin (see page 27-35).

In light of the teaching in the Chamberlain Provisional, it is clear that the Chamberlain Patent Pub anticipates exemplary Xiao Claim 1 under 35 USC 102(e). As such, allowed Xiao claims similar to or identical to this claim (that include the 4 rod repeats) should be withdrawn from issue as unpatentable over the Chamberlain Pat Pub.

iii. Chamberlain Patent Pub Anticipates All Exemplary Claims that Include 4 Repeats Under 35 USC 102(e)

The remaining exemplary claims from the Xiao Patent Pub that include 4 rod repeats are also anticipated by the Chamberlain Patent Pub based on the disclosure in the Chamberlain Provisional. A claim by claim analysis is provided below, showing where support for each element is found in the Chamberlain Provisional. It is noted that the support provided in the Chamberlain Provisional is merely exemplary and not intended to be comprehensive (i.e. additional support for each element may also be found throughout the remainder of the Provisional Application):

Claim 2: requires that the mini-dystrophin further comprises at least 3 amino acids of a C-terminal domain of the dystrophin gene. **Chamberlain Provisional** - Figure 27 teaches the use of "CT" (C-terminal domain) as part of the 4 rod repeat constructs. C-terminal domains are also taught on page 8, lines 8-18; and SEQ ID NO:36.

Claim 3: essentially identical to Claim 1, but requires the nucleic acid sequence to be less than 5,000 nucleotides in length. **Chamberlain Provisional** - Page 4 teaches that the nucleic acid "is less than 5 kilo-bases in length" and Example 6 describes a four rod repeat construct, deleted for exons 71-78, with a total size of 4.7kb.

Claim 4: requires that the mini-dystrophin further comprise an H2 or H3. **Chamberlain Provisional** - Figure 27 teaches R2-R21-H3-Micro-Dys 3 that contains H1, H4 and further contains H3.

Claim 5: requires 4 rod repeats specifically. **Chamberlain Provisional** - See discussion above of extensive teaching of 4 rod repeats in Chamberlain Provisional.

Claim 11: nucleic acid sequence consists of SEQ ID NO:10, which is construct Δ3531 containing 4 rod repeats (R1, R22, R23, and R24; see Figure 1 of Xiao Pat Pub). **Chamberlain Provisional** - teaches construct R2-R21-Micro-Dys 2 that contains the exact 4 repeats (i.e R1, R22, R23, and R24), as well as the same hinges (H1 and H4), and the use of an N-terminal, cystein-rich, and C-terminal domains. The sequence for Micro-Dys 2 is provided in Figure 13, which is SEQ ID NO:40.

Claim 12: nucleic acid sequence consists of SEQ ID NO:12, which is construct Δ3510 containing 4 rod repeats (R1, R2, R23, and R24; see Figure 1 of Xiao Pat Pub). **Chamberlain Provisional** - teaches construct R2-R21-Micro-Dys 2 that contains the 3 of the four repeats in SEQ ID NO:12 (Micro-Dys 2 uses R22 instead of R2) Micro-Dys 2 uses the same hinges (H1 and H4), and the use of an N-terminal, cystein-rich, and C-terminal domains. The sequence for Micro-Dys 2 is provided in Figure 13, which is SEQ ID NO:40.

Claim 13: nucleic acid sequence consists of SEQ ID NO:14, which is construct Δ3447 containing 4 rod repeats (R1, R2, R3, and R24; see Figure 1 of Xiao Pat Pub). **Chamberlain Provisional** - teaches construct R4-R23-Micro-Dys 1 that contains the exact 4 repeats (i.e. R1, R2, R3, and R24), as well as the same hinges (H1 and H4), and the use of an N-terminal, cystein-rich, and C-terminal domains. The sequence for Micro-Dys 1 is provided in Figure 11, which is SEQ ID NO:39.

Claim 14: requires the nucleic acid sequence of Claim 1 linked to expression control elements in an AAV. **Chamberlain Provisional** - teaches the use of AAV vectors at pages 40-41 and Example 6, as well as the use of expression control elements at pages 7, lines 2-6, and pages 36-38, and Figures 36-37.

Claim 15: limits claim 14 by requiring MCK or CMV promoters. **Chamberlain Provisional** - teaches the use of MCK promoters at pages 7, lines 2-6, and pages 36-38, and Figures 36-37.

Claim 18: Method of using nucleic acid of Claim 1 attached to an expression control element in order to treat Duchenne muscular dystrophy (DMD) or Becker muscular dystrophy (BMD). **Chamberlain Provisional** - Treating a subject with DMD or BMD is taught on page 8, lines 3-7 and Example 6.

Claim 19: Limits Claim 18 by requiring the use of AAV. **Chamberlain Provisional** - Example 6 which describes the use of a 4-repeat construct with AAV to treat DMD or BMD.

Claim 20: Limits Claim 18 by requiring the use of a retrovirus. **Chamberlain Provisional** - teaches the use of retroviruses on pages 41-42 under the heading "6. Retroviruses."

Claim 21: Essentially repeats Claim 18, but requires that the nucleic acid sequence be less than 5,000 bases. **Chamberlain Provisional** - Again, Example 6 describes the treatment of DMD or BMD with a construct that is 4.7 kb.

Claim 22: Further limits Claim 21 by requiring the use of AAV. **Chamberlain Provisional** - Example 6 uses AAV.

Claim 23: Further limits Claim 21 by requiring the use of a retrovirus. **Chamberlain Provisional** - teaches the use of retroviruses on pages 41-42 under the heading "6. Retroviruses."

Claim 26: Dependent on Claim 11, further requiring an expression control element. **Chamberlain Provisional** - see support for Claim 11 and discussion of support for control elements above.

Claim 27: Dependent on Claim 12, further requiring an expression control element. **Chamberlain Provisional** - see support for Claim 12 and discussion of support for control elements above.

Claim 28: Dependent on Claim 13, further requiring an expression control element. **Chamberlain Provisional** - see support for Claim 13 and discussion of support for control elements above.

In light of the teaching in the Chamberlain Provisional described above, it is clear that the Chamberlain Patent Pub anticipates all of the exemplary Xiao Claims that include a 4 repeat mini-dystrophin under 35 USC 102(e). As such, allowed Xiao claims similar to or identical to these claim (that include the 4 rod repeats) should be withdrawn from issue as unpatentable over the Chamberlain Pat Pub.

III. Withdrawal from Issuance Under 37 CFR 1.313

To the extent any of the Xiao claims that have been allowed include 4 repeat mini-dystrophins, these claims must be withdrawn from allowance. The Patent Office regulations specifically contemplate that an application may withdrawn after it has been allowed. In particular, 37 CFR 1.313, entitled "Withdrawal from issue" indicates that an Application "may be withdrawn from issue for further action at the initiative of the Office or upon petition by the Applicant." The wording of section 1.313(a) makes it clear that the Examiner can withdraw the

Application for any reason. The wording of section 1.313(b) makes it clear that the Examiner can still withdraw the Application after the issue fee has been paid for a number of reasons, including: (1) Unpatentability of one or more claims; or (2) Interference.

While it is unknown by Protestors whether the issue fee has been paid (which is unlikely given that the Notice of Allowance was only mailed about 12 days ago), it is not relevant because this case should be withdrawn based on the unpatentability of any claims that encompass 4 repeat mini-dystrophins in light of the Chamberlain Pat Pub. It is also noted that, under MPEP 1308, even if the issue fee has been paid, an allowed Application:

> "may be removed from the Office of Patent Publication, without it being withdrawn from issue under 37 CFR 1.313(b) to permit the Examiner to consider an information disclosure statement or whether one or more claims are unpatentable."

Therefore, at a minimum, Protestors urge the Examiner to request that the file be returned to the Examiner to consider the patentability of the claims in light of the Chamberlain Pat Pub, in the interest of not issuing an invalid patent. Furthermore, Protestors submit that any claims that encompass 4 rod repeats must cause the application to be withdrawn from issuance such that the claims can be rejected in light of the Chamberlain Patent Pub.

Finally, as noted above, a second reason an Application may be withdrawn from issue (even if the issue fee has been paid) is for an Interference. Protestors are currently prosecuting U.S. Application Ser. No. 10/149,736, which published as the Chamberlain Pat Pub. This Application is currently in active prosecution, with claims directed specifically to nucleic acid sequences encoding mini-dystrophins with 4 rod repeats. Prior to declaring an Interference, since the Chamberlain Application would be the senior party with respect to 4 repeat constructs, Protestors submit that the Examiner should reject any Xiao claims that include 4 repeat constructs. Only IF Xiao is able to demonstrate invention prior to the Chamberlain Provisional filing date would an Interference be possible.

Conclusion

To the extent any of the allowed Xiao claims encompass 4 rod repeat mini-dystrophins, these claims must be withdrawn from issuance and rejected in light of the Chamberlain Patent Pub. Protestors submit that it would be unfair to the public to allow any 4 repeat encompassing Xiao claims to issue as these claims would be invalid in light of the Chamberlain Pat Pub.

Applicants note for the Examiner's convenience that a copy of this communication will be sent to Xiao's representative at the correspondence address listed on the cover of the Xiao Pat Pub, as well as to the Examiner handling the pending Chamberlain Application (with the next communication in that case).

Dated: May 10, 2004

Jason R. Bond
Registration No. 45,439
MEDLEN & CARROLL, LLP
440 Science Drive, Suite 203
Madison, Wisconsin 53711
608-218-6900

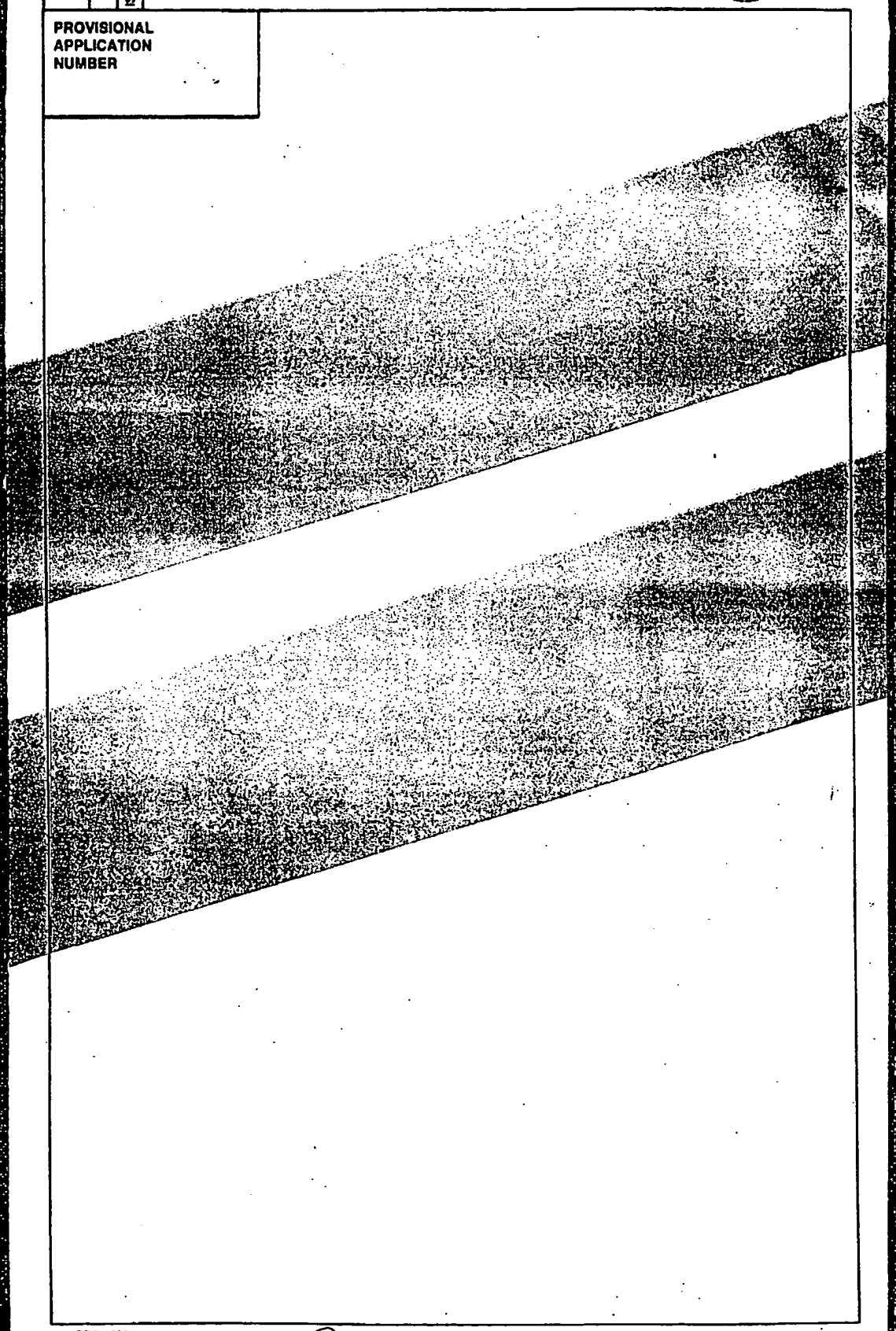
04/28/00

ISSUE CLASSIFICATION	Class	Subclass

[redacted]

C

PROVISIONAL APPLICATION NUMBER



Form PTO-1525
(Rev. 6/99)

(FACE)

PATENT APPLICATION

60200777

APPROVED FOR LICENSE

INITIALS

CONTENTS

Date Entered or Counted		Date Received or Mailed
	1. Application papers	
	2. Request for Licence 3/03/09	
	3.	
	4.	
	5.	
	6.	
	7.	
	8.	
	9.	
	10.	
	11.	
	12.	
	13.	
	14.	
	15.	
	16.	
	17.	
	18.	
	19.	
	20.	
	21.	
	22.	
	23.	
	24.	
	25.	
	26.	
	27.	
	28.	
	29.	
	30.	
	31.	
	32.	

(LEFT OUTSIDE)

POSITION	INITIALS	ID NO.	DATE
FEE DETERMINATION			
O.I.P.E. CLASSIFIER			
FORMALITY REVIEW	MT	166548	6/29/00
RESPONSE FORMALITY REVIEW			

(LEFT INSIDE)

PATENT APPLICATION SERIAL NO. _____

U.S. DEPARTMENT OF COMMERCE
PATENT AND TRADEMARK OFFICE
FEE RECORD SHEET

05/04/2000 TLUU11 00000122 011944 60200777
01 FC:214 75.00 CH

PTO-1556
(5/87)

Bib Data Sheet

UNITED STATES DEPARTMENT OF COMMERCE
Patent and Trademark Office
Address: COMMISSIONER OF PATENTS AND TRADEMARKS
Washington, D.C. 20231

SERIAL NUMBER	FILING DATE	CLASS	GROUP ART UNIT	ATTORNEY DOCKET NO.
60/200,777	04/28/2000 RULE _	-	-	4268

APPLICANTS
Xiao Xiao, Wexford, PA ;

** CONTINUING DATA **********

** FOREIGN APPLICATIONS **********

IF REQUIRED, FOREIGN FILING LICENSE GRANTED ** SMALL ENTITY **
** 06/29/2000

Foreign Priority claimed	STATE OR COUNTRY	SHEETS DRAWING	TOTAL CLAIMS	INDEPENDENT CLAIMS
☐ yes ☐ no	PA	4	-	-
35 USC 119 (a-d) conditions met ☐ yes ☐ no ☐ Met after Allowance				
Verified and Acknowledged Examiner's Signature Initials				

ADDRESS

Anderson Kill and Olick P C
1251 Avenue of the Americas
New York ,NY 10020-1182

TITLE

Adeno-associated viral vectors carrying novel human mini-dystrophin genes

FILING FEE RECEIVED	FEES: Authority has been given in Paper	
75	No. _____ to charge/credit DEPOSIT ACCOUNT No. _____ for following:	☐ All Fees ☐ 1.16 Fees (Filing) ☐ 1.17 Fees (Processing Ext. of time) ☐ 1.18 Fees (Issue) ☐ Other _____ ☐ Credit

Docket No. 4268

PROVISIONAL PATENT APPLICATION FOR:

Adeno-associated viral vectors carrying novel human mini-dystrophin genes

Mailing Address: Xiao Xiao, Ph.D.

2433 Dogwood Drive

Wexford, PA 15090

(724) 934-9798

(412) 648-9610 (fax)

xiaox@pitt.edu

Blue dye positive myofibers were observed under the fluorescent microscope with Rhodamine filters.

Abstract

The present invention provides a series of novel mini-dystrophin genes that retain the essential biological functions. The expression of the mini-dystrophin genes are under the control of a muscle-specific promoter or a non-muscle-specific promoter along with a small polyadenylation signal. The entire gene expression cassettes can be readily packaged into adeno-associated virus (AAV) vectors. Moreover, the present invention provides a method of treatment for Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy BMD), using the novel AAV vectors containing the mini-genes. These mini-dystrophin genes have been demonstrated by the inventor in a DMD mouse model *mdx* to be able to alleviate muscular dystrophic pathology and to result in normal myofiber morphology, histology and cell membrane integrity. Finally, the present invention further defines the minimal functional domains of dystrophin and provides ways to optimize and create new versions of mini-dystrophin genes.

TECHNICAL FIELD

The present invention provides a series of novel mini-dystrophin genes that retain the essential biological functions. The expression of the mini-dystrophin genes are under the control of a muscle-specific promoter or a non-muscle-specific promter along with a small polyadenylation signal. The entire gene expression cassettes can be readily packaged into adeno-associated virus (AAV) vectors. Moreover, the present invention provides an method of treatment for Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy BMD) using the novel AAV vectors containing the mini-genes, which are demonstrated by the inventor in a DMD mouse model *mdx* to alleviate muscular dystrophic pathology and to result in normal myofiber morphology, histology and cell membrane integrity. Finally, the present invention further defines the minimal functional domains of dystrophin and provides ways to optimize and create new versions of mini-dystrophin genes.

BACKGROUND OF INVENTION

Duchenne muscular dystrophy (DMD) is the most common and lethal genetic muscle disorder, caused by recessive mutations in dystrophin gene. [Koenig, M. et al. *Cell* 50, 509-517 (1987)]. One of every 3500 males suffers from DMD, yet no treatment is available. Genetic therapeutic approaches using primarily myoblast transplantation or adenovirus-mediated gene transfer, have met with little success. [Partridge, T. A. et al., *Nature* 337, 176-179 (1989); Acsadi, G. et al. *Nature* 352, 815-818 (1991); Ragot, T. *et al. Nature* 361, 647-650 (1993); and Gussoni, E., et al., *Nature Med.* 3, 970-977 (1997)]. Adeno-associated virus (AAV) vectors, although proven superior for muscle gene transfer, [Xiao, X. et al., *Journal of Virology* 70, 8098-8108 (1996)] are too small (5 kb) to package the dystrophin gene (14 kb cDNA).

Gene delivery is an important method for the treatment of acquired and inherited diseases. A number of viral based systems are being developed for gene transfer purposes. In particular, retroviruses are currently the most widely used viral vector system for gene delivery. Although retroviral systems are popular, they suffer from several drawbacks. Especially, retroviral particles are relatively labile and hence unstable. Therefore, purification of recombinant viruses can lead to significant loss in titer. Moreover, retroviruses have a limited host range and cannot integrate into nonreplicating cells. Accordingly, cells do not normally divide, such as mature muscle myofibers and neurons, or cells which replicate slowly, cannot be genetically altered using retroviral vectors unless stimulated to divide before infection. Furthermore, retroviruses are known to cause disease in certain animals, including humans, and thus pose a significant health risk to the subject transfected with a recombinant virus.

Adenovirus based systems have been developed for gene delivery in an attempt to overcome these problems. Human adenoviruses are double-stranded DNA viruses which enter cells by receptor-mediated endocytosis. These viruses are particularly well suited for gene transfer because they are easy to grow and manipulate and they exhibit a broad host range in vivo and in vitro. Despite these advantages, adenovirus vectors suffer from several drawbacks. For example, adenovirus vectors express proteins transiently because the transferred gene does not integrate into the chromosome of the target cell. Hence, as the cells divide, the transferred gene is lost. In this regard, such vectors are ineffective for long term gene therapy. Furthermore, adenovirus vectors express viral proteins that may elicit an immune response which may decrease the life of the transduced cell. This immune response may preclude subsequent treatments because humoral and/or T cell responses. Finally, Ad vectors can not efficiently infect mature muscle due to the inability to bypass the barrier of extracellular matrix.

Adeno-associatee virus (AAV), the only non-pathogenic viral vector currently available, has been successfully used to establish efficient and long-term gene expression in both dividing and non-dividing cells *in vivo* without significant immune response or toxicity [Samulski, R. J. et al *Development of Human Gene Therapy* 131-172 ,Cold Spring Harbor Laboratory Press, Cold Spring Harbor, (1999)]. Unlike other viral and non-viral vectors, AAV readily bypasses extracellular barriers in muscle due to its small particle size (20 nm), and it successfully transduces myofibers of various maturity. Currently, AAV vectors offer the best gene transfer efficiency and longevity among all viral and non-viral vectors tested for gene therapy in muscle tissues. This unparalleled efficiency and safety of the vector system have led to an increasing interest in AAV-mediated gene therapy for genetic muscle disorders as well as for metabolic diseases involving genes of smaller size (< 4.5 kb). [Kessler, P. D. *et al. Proceedings of the National Academy of Sciences of the United States of America* 93, 14082-14087 (1996); Song, S. *et al. Proc Natl Acad Sci U S A* 95, 14384-14388 (1998); Herzog, R. W. *et al. Nat Med* 5, 56-63 (1999); and Xiao, X. *et al. J. Virol.* 74, 1436-1442 (2000)].

Previous attempts to generate mini-genes that were shorter than 1/2 of the full length dystrophin failed to preserve the essential protective functions. [Yuasa, K. *et al., FEBS Lett* 425, 329-336 (1998)]. Although the mini-genes contained both intact N- and C-terminal domains and 1 to 3 central rod repeats, they were functionally similar to a C-terminal dystrophin construct (Dp71), [Cox, G. A. et al., *Nat Genet* 8, 333-339 (1994); Greenberg, D. S. et al., *Nat Genet* 8, 340-344 (1994)], and thus sufficient to restore DAP complexes but insufficient to protect muscle from dystrophic pathology. However, the mini-dystrophin genes reported in the present invention accommodated at least 5 rod repeats (R1, R2, R22, R23 and R24) and two hinges (H1 and H4). Therefore we hypothesized that the length of the central rod domain is the most critical factor, based on the fact that a primary role of dystrophin is to crosslink the myofiber cytoskelton and plasma membrane and stabilize the structure during muscle contraction. If the dystrophin is too short to span the sliding distance between the cytoskeleton and plasma membrane during muscle contraction, the crosslink will be disrupted and the muscle membrane will become

unstable and prone to mechanical damages. To accommodate as many rod units in the central domain without exceeding the AAV vector packaging limit, we have for the first time deleted the entire C-terminus (819 bp) without sacrificing the primary functions of dystrophin. Our results indicate that 5 rods and 2 hinges provide sufficient length and flexibility for the central domain.

The most importance is that AAV is the best vector system currently available for muscle-based gene therapy. However, until this report AAV's utility has been precluded for DMD, the most common and lethal muscle disorders. Previously other viral and nonviral vectors, [Acsadi, G. *et al. Nature* **352**, 815-818 (1991); Ragot, T. *et al. Nature* **361**, 647-650 (1993)] as well as myoblast transplantation [Partridge, T. A. et al., *Nature* **337**, 176-179 (1989); Gussoni, E. et al., *Narure Med.* **3**, 970-977 (1997)] have been explored for DMD with limited success. Recent studies using stem-cell transplantation have offered a new hope for cell therapeutics of DMD. [Gussoni, E. *et al. Nature* **401**, 390-394 (1999)]. The novel functional dystrophin genes reported here should also find their utilities in the stem-cell therapy after *ex vivo* gene transfer. Nevertheless, the primary advantage of AAV vector is its direct *in vivo* gene delivery such as intramuscular injections, or *in vivo* vector delivery through blood circulation [Greelish, J. P. *et al. Nat Med* **5**, 439-443 (1999)] . Finally, using the AAV vector rather than the traditional transgenic mouse technology, we have provided a more convenient and less time-consuming method to further discern the dystrophin functional domains *in vivo* and to optimize the mini-genes for DMD gene therapy.

SUMMARY OF THE INVENTION

The present invention provides the dystrophin gene which can be successfully reduced to approximately one third (1/3) of its 11 kb full-length coding sequence, without compromising essential functions in protecting muscles from dystrophic phenotypes. Moreover, the present invention provides AAV vectors carrying the mini-genes and capable of mediating efficient and stable correction of both biochemical and physiological defects in a major muscle group of a DMD animal model. Furthermore, the expression of the mini-dystrophin genes are controlled either by a muscle-specific promoter, or by a non-muscle-specific promoter along with a small polyadenylation signal. New development in systemic delivery of AAV vectors through the blood circulation should enable more widespread gene transfer in large groups of muscle for DMD gene therapy. Furthermore, the present invention provides a method that is more convenient and less time-consuming to discern the dystrophin functional domains *in vivo* and to optimize the mini-genes for DMD gene therapy.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1. Construction of highly truncated mini-dystrophin genes.

Dystrophin has four major domains: the N-terminal domain (N), the cysteine-rich domain (CR), the C-terminal domain (CT) and the central rod domain, which contains 24 rod repeats (R) and 4 hinges (H). The mini-dystrophin genes were constructed by deleting a large portion of the central rods and hinges and nearly the entire CT domain (except the last 5 amino acids). The mini-dystrophin genes were subsequently cloned between an MCK (muscle-specific creatine kinase) promoter, or a CMV promoter, along with a small polyA sequence in AAV vectors.

Figure 2. Immunofluorescent (IF) analysis of the dystrophin and dystrophin-associated protein complexes in gastrocnemius muscle.

a. Cryosections of mdx muscle at 3-months after treatment with construct MCK-3849 (A) or MCK-3990 (B) were IF stained with an antibody against dystrophin (Dys3, green color) and counter-stained for cell nuclei with DAPI (blue color). Photos were taken with a 4X microscope lens. Note that widespread mini-gene expression and healthy peripheral nucleation are evident.

b. **b.** Cryosections of muscles from 15-week old normal C57/B10 mice, or mdx mice treated either with vector MCK-3849, MCK-3990 or MCK-4173, or untreated mdx mice, were IF stained with antibodies against dystrophin (1^{st} row), or against α-sarcoglycan (2^{nd} row), β-sarcoglycan (3^{rd} row) and γ-sarcoglycan (4^{th} row). Cryosections stained with anti-dystrophin antibody (1^{st} row) were also counter-stained with DAPI (blue color) for cell nuclei. Photos were taken with a 20X microscope lens.

Figure 3. In vivo myofiber plasma membrane integrity test.

At 15 hours after intravenous injection of Evans Blue dye, the gastrocnemius muscle either from 15-week old normal C57/B10 mice, from mdx mice treated with AAV vectors containing mini-genes Δ3849, Δ3990 or Δ4173, or from the untreated mdx mice were collected and cryosectioned. Dystrophin or mini-dystrophin expression was visualized by immunofluorescent staining (1st column, green color). The leaky myofibers were visualized by the uptake of Evans Blue dye showing red fluorescence (2nd column). Note the mutual exclusivity between dystrophin expression and Evans Blue dye uptake, when the two images were superimposed (3rd column). Photos were taken with a 10X microscope lens.

DNA Sequence ID No 1 to No 9:

Seq. ID No 1: Mini-dystrophin Δ4173 It is a mini-dystrophin construct.

AttttcaCCatgGtttggtgggaagaagtagaggactgttatgaaagagaagatgttcaaaagaaaacattcacaaaatg

ggtaaatgcacaattttctaagtttgggaagcagcatattgagaacctcttcagtgacctacaggatgggaggcgcctcc

tagacctcctcgaaggcctgacagggcaaaaactgccaaaagaaaaaggatccacaagagttcatgccctgaacaatgtc

aacaaggcactgcgggttttgcagaacaataatgttgatttagtgaatattggaagtactgacatcgtagatggaaatca

taaactgactcttggtttgatttggaatataatcctccactggcaggtcaaaaatgtaatgaaaaatatcatggctggat

tgcaacaaaccaacagtgaaaagattctcctgagctgggtccgacaatcaactcgtaattatccacaggttaatgtaatc

aacttcaccaccagctggtctgatggcctggctttgaatgctctcatccatagtcataggccagacctatttgactggaa

tagtgtggtttgccagcagtcagccacacaacgactggaacatgcattcaacatcgccagatatcaattaggcatagaga

aactactcgatcctgaagatgttgataccacctatccagataagaagtccatcttaatgtacatcacatcactcttccaa

gttttgcctcaacaagtgagcattgaagccatccaggaagtggaaatgttgccaaggccacctaaagtgactaaagaaga

acattttcagttacatcatcaaatgcactattctcaacagatcacggtcagtctagcacagggatatgagagaacttctt

cccctaagcctcgattcaagagctatgcctacacacaggctgcttatgtcaccacctctgaccctacacggagcccattt

ccttcacagcatttggaagctcctgaagacaagtcatttggcagttcattgatggagagtgaagtaaacctggaccgtta

tcaaacagctttagaagaagtattatcgtggcttctttctgctgaggacacattgcaagcacaaggagagatttctaatg

atgtggaagtggtgaaagaccagtttcatactcatgaggggtacatgatggatttgacagcccatcagggccgggttggt

aatattctacaattgggaagtaagctgattggaacaggaaaattatcagaagatgaagaaactgaagtacaagagcagat

gaatctcctaaattcaagatgggaatgcctcagggtagctagcatggaaaaacaaagcaatttacatagagttttaatgg

atctccagaatcagaaactgaaagagttgaatgactggctaacaaaaacagaagaaagaacaaggaaaatggaggaagag

cctcttggacctgatcttgaagacctaaaacgccaagtacaacaacataaggtgcttcaagaagatctagaacaagaaca

agtcagggtcaattctctcactcacatggtggtggtagttgatgaatctagtggagatcacgcaactgctgctttggaag

aacaacttaaggtattgggagatcgatgggcaaacatctgtagatggacagaagaccgctgggttcttttacaagacatc

cttctcaaatggcaacgtcttactgaagaacagtgcctttttagtgcatggctttcagaaaaagaagatgcagtgaacaa

gattcacacaactggctttaaagatcaaaatgaaatgttatcaagtcttcaaaaactggccgttttaaaagcggatctag

aaaagaaaaagcaatccatgggcaaactgtattcactcaaacaagatcttctttcaacactgaagaataagtcagtgacc

cagaagacggaagcatggctggataactttgcccggtgttgggataatttagtccaaaaacttgaaaagagtacagcaca

gactcatagattactgcaacagttccccctggacctggaaaagtttcttgcctggcttacagaagctgaaacaactgcca

atgtcctacaggatgctacccgtaaggaaaggctcctagaagactccaagggagtaaaagagctgatgaaacaatggcaa

gacctccaaggtgaaattgaagctcacacagatgtttatcacaacctggatgaaaacagccaaaaaatcctgagatccct

ggaaggttccgatgatgcagtcctgttacaaagacgtttggataacatgaacttcaagtggagtgaacttcggaaaaagt

ctctcaacattaggtcccatttggaagccagttctgaccagtggaagcgtctgcacctttctctgcaggaacttctggtg

tggctacagctgaaagatgatgaattaagccggcaggcacctattggaggcgactttccagcagttcagaagcagaacga

tgtacatagggccttcaagagggaattgaaaactaaagaacctgtaatcatgagtactcttgagactgtacgaatatttc

tgacagagcagcctttggaaggactagagaaactctaccaggagcccagagagctgcctcctgaggagagagcccagaat

gtcactcggcttctacgaaagcaggctgaggaggtcaatactgagtgggaaaaattgaacctgcactccgctgactggca

gagaaaaatagatgagacccttgaaagactccaggaacttcaagaggccacggatgagctggacctcaagctgcgccaag

ctgaggtgatcaagggatcctggcagcccgtgggcgatctcctcattgactctctccaagatcacctcgagaaagtcaag

gcacttcgaggagaaattgcgcctctgaaagagaacgtgagccacgtcaatgaccttgctcgccagcttaccactttggg

cattcagctctcaccgtataacctcagcactctggaagacctgaacaccagatggaagcttctgcaggtggccgtcgagg

accgagtcaggcagctgcatgaagcccacagggactttggtccagcatctcagcactttctttccacgtctgtccagggt

ccctgggagagagccatctcgccaaacaaagtgccctactatatcaaccacgagactcaaacaacttgctgggaccatcc

caaaatgacagagctctaccagtctttagctgacctgaataatgtcagattctcagcttataggactgccatgaaactcc

gaagactgcagaaggccctttgcttggatctcttgagcctgtcagctgcatgtgatgccttggaccagcacaacctcaag

caaaatgaccagcccatggatatcctgcagattattaattgtttgaccactatttatgaccgcctggagcaagagcacaa

caatttggtcaacgtccctctctgcgtggatatgtgtctgaactggctgctgaatgtttatgatacgggacgaacaggga

ggatccgtgtcctgtcttttaaaactggcatcatttccctgtgtaaagcacatttggaagacaagtacagataccttttc

aagcaagtggcaagttcaacaggattttgtgaccagcgcaggctgggcctccttctgcatgattctatccaaattccaag

acagttgggtgaagttgcatcctttgggggcagtaacattgagccaagtgtccggagctgcttccaatttgctaataata

agccagagatcgaagcggccctcttcctagactggatgagactggaaccccagtccatggtgtggctgcccgtcctgcac

agagtggctgctgcagaaactgccaagcatcaggccaaatgtaacatctgcaaagagtgtccaatcattggattcaggta

caggagtctaaagcactttaattatgacatctgccaaagctgctttttttctggtcgagttgcaaaaggccataaaatgc

actatcccatggtggaatattgcactccgactacatcaggagaagatgttcgagactttgccaaggtactaaaaaacaaa

tttcgaaccaaaaggtattttgcgaagcatccccgaatgggctacctgccagtgcagactgtcttagaggggacaacat

ggaaactcccgacacaatgtag

Seq. ID No 2; Mini-dystrophin Δ3990. It is a mini-dystrophin construct.

AttttcaCCatgGtttggtgggaagaagtagaggactgttatgaaagagaagatgttcaaaagaaaacattcacaaaatg

ggtaaatgcacaattttctaagtttgggaagcagcatattgagaacctcttcagtgacctacaggatgggaggcgcctcc

tagacctcctcgaaggcctgacagggcaaaaactgccaaaagaaaaaggatccacaagagttcatgccctgaacaatgtc

aacaaggcactgcgggttttgcagaacaataatgttgatttagtgaatattggaagtactgacatcgtagatggaaatca

taaactgactcttggtttgatttggaatataatcctccactggcaggtcaaaaatgtaatgaaaaatatcatggctggat

tgcaacaaaccaacagtgaaaagattctcctgagctgggtccgacaatcaactcgtaattatccacaggttaatgtaatc

aacttcaccaccagctggtctgatggcctggctttgaatgctctcatccatagtcataggccagacctatttgactggaa

tagtgtggtttgccagcagtcagccacacaacgactggaacatgcattcaacatcgccagatatcaattaggcatagaga

aactactcgatcctgaagatgttgataccacctatccagataagaagtccatcttaatgtacatcacatcactcttccaa

gttttgcctcaacaagtgagcattgaagccatccaggaagtggaaatgttgccaaggccacctaaagtgactaaagaaga

acattttcagttacatcatcaaatgcactattctcaacagatcacggtcagtctagcacagggatatgagagaacttctt

cccctaagcctcgattcaagagctatgcctacacacaggctgcttatgtcaccacctctgaccctacacggagcccattt

ccttcacagcatttggaagctcctgaagacaagtcatttggcagttcattgatggagagtgaagtaaacctggaccgtta

tcaaacagctttagaagaagtattatcgtggcttctttctgctgaggacacattgcaagcacaaggagagatttctaatg

atgtggaagtggtgaaagaccagtttcatactcatgaggggtacatgatggatttgacagcccatcagggccgggttggt

aatattctacaattgggaagtaagctgattggaacaggaaaattatcagaagatgaagaaactgaagtacaagagcagat

gaatctcctaaattcaagatgggaatgcctcagggtagctagcatggaaaaacaaagcaatttacatagagtttaatgg

atctccagaatcagaaactgaaagagttgaatgactggctaacaaaaacagaagaaagaacaaggaaaatggaggaagag

cctcttggacctgatcttgaagacctaaaacgccaagtacaacaacataaggtgcttcaagaagatctagaacaagaaca

agtcagggtcaattctctcactcacatggtggtggtagttgatgaatctagtggagatcacgcaactgctgctttggaag

aacaacttaaggtattgggagatcgatgggcaaacatctgtagatggacagaagaccgctgggttcttttacaagaccag

cctgacctagctcctggactgaccactattggagcctctcctactcagactgttactctggtgacacaacctgtggttac

taaggaaactgccatctccaaactagaaatgccatcttccttgatgttggaggtacctactcatagattactgcaacagt

tccccctggacctggaaaagtttcttgcctggcttacagaagctgaaacaactgccaatgtcctacaggatgctacccgt

aaggaaaggctcctagaagactccaagggagtaaaagagctgatgaaacaatggcaagacctccaaggtgaaattgaagc

tcacacagatgtttatcacaacctggatgaaaacagccaaaaaatcctgagatccctggaaggttccgatgatgcagtcc

tgttacaaagacgtttggataacatgaacttcaagtggagtgaacttcggaaaaagtctctcaacattaggtcccatttg

gaagccagttctgaccagtggaagcgtctgcacctttctctgcaggaacttctggtgtggctacagctgaaagatgatga

attaagccggcaggcacctattggaggcgactttccagcagttcagaagcagaacgatgtacatagggccttcaagaggg

aattgaaaactaaagaacctgtaatcatgagtactcttgagactgtacgaatatttctgacagagcagcctttggaagga

ctagagaaactctaccaggagcccagagagctgcctcctgaggagagagcccagaatgtcactcggcttctacgaaagca

ggctgaggaggtcaatactgagtgggaaaaattgaacctgcactccgctgactggcagagaaaaatagatgagacccttg

aaagactccaggaacttcaagaggccacggatgagctggacctcaagctgcgccaagctgaggtgatcaagggatcctgg

cagcccgtgggcgatctcctcattgactctctccaagatcacctcgagaaagtcaaggcacttcgaggagaaattgcgcc

tctgaaagagaacgtgagccacgtcaatgaccttgctcgccagcttaccactttgggcattcagctctcaccgtataacc

tcagcactctggaagacctgaacaccagatggaagcttctgcaggtggccgtcgaggaccgagtcaggcagctgcatgaa

gcccacagggactttggtccagcatctcagcacttctttccacgtctgtccagggtccctgggagagagccatctcgcc

aaacaaagtgccctactatatcaaccacgagactcaaacaacttgctgggaccatcccaaaatgacagagctctaccagt

ctttagctgacctgaataatgtcagattctcagcttataggactgccatgaaactccgaagactgcagaaggccctttgc

ttggatctcttgagcctgtcagctgcatgtgatgccttggaccagcacaacctcaagcaaaatgaccagcccatggatat

cctgcagattattaattgtttgaccactatttatgaccgcctggagcaagagcacaacaatttggtcaacgtccctctct

gcgtggatatgtgtctgaactggctgctgaatgtttatgatacgggacgaacagggaggatccgtgtcctgtcttttaaa

actggcatcatttccctgtgtaaagcacatttggaagacaagtacagataccttttcaagcaagtggcaagttcaacagg

attttgtgaccagcgcaggctgggcctccttctgcatgattctatccaaattccaagacagttgggtgaagttgcatcct

ttgggggcagtaacattgagccaagtgtccggagctgcttccaatttgctaataataagccagagatcgaagcggccctc

ttcctagactggatgagactggaaccccagtccatggtgtggctgcccgtcctgcacagagtggctgctgcagaaactgc

caagcatcaggccaaatgtaacatctgcaaagagtgtccaatcattggattcaggtacaggagtctaaagcactttaatt

atgacatctgccaaagctgctttttttctggtcgagttgcaaaaggccataaaatgcactatcccatggtggaatattgc

actccgactacatcaggagaagatgttcgagactttgccaaggtactaaaaaacaaatttcgaaccaaaaggtattttgc

gaagcatccccgaatgggctacctgccagtgcagactgtcttagaggggacaacatggaaactcccgacacaatgtag

Seq. ID No 3: Mini-dystrophin Δ3849. It is a mini-dystrophin construct.

AttttcaCCatgGtttggtgggaagaagtagaggactgttatgaaagagaagatgttcaaaagaaaacattcacaaaatg

ggtaaatgcacaattttctaagtttgggaagcagcatattgagaacctcttcagtgacctacaggatgggaggcgcctcc

tagacctcctcgaaggcctgacagggcaaaaactgccaaaagaaaaaggatccacaagagttcatgccctgaacaatgtc

aacaaggcactgcgggtttgcagaacaataatgttgatttagtgaatattggaagtactgacatcgtagatggaaatca

taaactgactcttggtttgatttggaatataatcctccactggcaggtcaaaaatgtaatgaaaaatatcatggctggat

tgcaacaaaccaacagtgaaaagattctcctgagctgggtccgacaatcaactcgtaattatccacaggttaatgtaatc

aacttcaccaccagctggtctgatggcctggctttgaatgctctcatccatagtcataggccagacctatttgactggaa

tagtgtggtttgccagcagtcagccacacaacgactggaacatgcattcaacatcgccagatatcaattaggcatagaga

aactactcgatcctgaagatgttgataccacctatccagataagaagtccatcttaatgtacatcacatcactcttccaa

gttttgcctcaacaagtgagcattgaagccatccaggaagtggaaatgttgccaaggccacctaaagtgactaaagaaga

acattttcagttacatcatcaaatgcactattctcaacagatcacggtcagtctagcacagggatatgagagaacttctt

cccctaagcctcgattcaagagctatgcctacacacaggctgcttatgtcaccacctctgaccctacacggagcccattt

ccttcacagcatttggaagctcctgaagacaagtcatttggcagttcattgatggagagtgaagtaaacctggaccgtta

tcaaacagctttagaagaagtattatcgtggcttctttctgctgaggacacattgcaagcacaaggagagatttctaatg

atgtggaagtggtgaaagaccagtttcatactcatgaggggtacatgatggatttgacagcccatcagggccgggttggt

aatattctacaattgggaagtaagctgattggaacaggaaaattatcagaagatgaagaaactgaagtacaagagcagat

gaatctcctaaattcaagatgggaatgcctcagggtagctagcatggaaaaacaaagcaatttacatagagttttaatgg

atctccagaatcagaaactgaaagagttgaatgactggctaacaaaaacagaagaaagaacaaggaaaatggaggaagag

cctcttggacctgatcttgaagacctaaaacgccaagtacaacaacataaggtgcttcaagaagatctagaacaagaaca

agtcagggtcaattctctcactcacatggtggtggtagttgatgaatctagtggagatcacgcaactgctgctttggaag

aacaacttaaggtattgggagatcgatgggcaaacatctgtagatggacagaagaccgctgggttcttttacaagacact

catagattactgcaacagttccccctggacctggaaaagtttcttgcctggcttacagaagctgaaacaactgccaatgt

cctacaggatgctacccgtaaggaaaggctcctagaagactccaagggagtaaaagagctgatgaaacaatggcaagacc

tccaaggtgaaattgaagctcacacagatgtttatcacaacctggatgaaaacagccaaaaaatcctgagatccctggaa

ggttccgatgatgcagtcctgttacaaagacgtttggataacatgaacttcaagtggagtgaacttcggaaaaagtctct

caacattaggtcccatttggaagccagttctgaccagtggaagcgtctgcacctttctctgcaggaacttctggtgtggc

tacagctgaaagatgatgaattaagccggcaggcacctattggaggcgactttccagcagttcagaagcagaacgatgta

catagggccttcaagagggaattgaaaactaaagaacctgtaatcatgagtactcttgagactgtacgaatatttctgac

agagcagcctttggaaggactagagaaactctaccaggagcccagagagctgcctcctgaggagagagcccagaatgtca

ctcggcttctacgaaagcaggctgaggaggtcaatactgagtgggaaaaattgaacctgcactccgctgactggcagaga

aaaatagatgagacccttgaaagactccaggaacttcaagaggccacggatgagctggacctcaagctgcgccaagctga

ggtgatcaagggatcctggcagcccgtgggcgatctcctcattgactctctccaagatcacctcgagaaagtcaaggcac

ttcgaggagaaattgcgcctctgaaagagaacgtgagccacgtcaatgaccttgctcgccagcttaccactttgggcatt

cagctctcaccgtataacctcagcactctggaagacctgaacaccagatggaagcttctgcaggtggccgtcgaggaccg

agtcaggcagctgcatgaagcccacagggactttggtccagcatctcagcactttcttccacgtctgtccagggtccct

gggagagagccatctcgccaaacaaagtgccctactatatcaaccacgagactcaaacaacttgctgggaccatcccaaa

atgacagagctctaccagtctttagctgacctgaataatgtcagattctcagcttataggactgccatgaaactccgaag

actgcagaaggccctttgcttggatctcttgagcctgtcagctgcatgtgatgccttggaccagcacaacctcaagcaaa

atgaccagcccatggatatcctgcagattattaattgtttgaccactatttatgaccgcctggagcaagagcacaacaat

ttggtcaacgtccctctctgcgtggatatgtgtctgaactggctgctgaatgtttatgatacgggacgaacagggaggat

ccgtgtcctgtcttttaaaactggcatcatttccctgtgtaaagcacatttggaagacaagtacagataccttttcaagc

aagtggcaagttcaacaggattttgtgaccagcgcaggctgggcctccttctgcatgattctatccaaattccaagacag

ttgggtgaagttgcatcctttgggggcagtaacattgagccaagtgtccggagctgcttccaatttgctaataataagcc

agagatcgaagcggccctcttcctagactggatgagactggaaccccagtccatggtgtggctgcccgtcctgcacagag

tggctgctgcagaaactgccaagcatcaggccaaatgtaacatctgcaaagagtgtccaatcattggattcaggtacagg

agtctaaagcactttaattatgacatctgccaaagctgctttttttctggtcgagttgcaaaaggccataaaatgcacta

tcccatggtggaatattgcactccgactacatcaggagaagatgttcgagactttgccaaggtactaaaaaacaaatttc

gaaccaaaaggtattttgcgaagcatccccgaatgggctacctgccagtgcagactgtcttagaggggacaacatggaa

actcccgacacaatgtag

Seq. ID No 4: AAV-MCK-Δ4173. It is an AAV vector containing a MCK promoter driving a mini-dystrophin gene connected to a small polyA signal.

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTC
GCCCGACGCCCGGGCTTTGCCCG

GGCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTA
GGGGTTCCTAGATCTGAATTCGag

cttgcatgcccactacgggtctaggctgcccatgtaaggaggcaaggcctggggacacccgagatgcctggttataatta

acccagacatgtggctgccccccccccccccaacacctgctgcctgagcctcacccccacccccggtgcctgggtcttaggc

tctgtacaccatggaggagaagctcgctctaaaaataaccctgtccctggtggatcccctgcatgcccaatcaaggctgt

gggggactgagggcaggctgtaacaggcttgggggccagggcttatacgtgcctgggactcccaaagtattactgttcca

tgttcccggcgaagggccagctgtcccccgccagctagactcagcacttagtttaggaaccagtgagcaagtcagcccctt

ggggcagcccatacaaggccatggggctgggcaagctgcacgcctgggtccggggtgggcacggtgcccgggcaacgagc

tgaaagctcatctgctctcaggggcccctccctggggacagcccctcctggctagtcacaccctgtaggctcctctatat

aacccaggggcacaggggctgcccccgggtcacTCGAAttttcaCCatgGtttggtgggaagaagtagaggactgttatg

aaagagaagatgttcaaaagaaaacattcacaaaatgggtaaatgcacaattttctaagtttgggaagcagcatattgag

aacctcttcagtgacctacaggatgggaggcgcctcctagacctcctcgaaggcctgacagggcaaaaactgccaaaaga

aaaaggatccacaagagttcatgccctgaacaatgtcaacaaggcactgcgggttttgcagaacaataatgttgatttag

tgaatattggaagtactgacatcgtagatggaaatcataaactgactcttggtttgatttggaatataatcctccactgg

caggtcaaaaatgtaatgaaaaatatcatggctggattgcaacaaaccaacagtgaaaagattctcctgagctgggtccg

acaatcaactcgtaattatccacaggttaatgtaatcaacttcaccaccagctggtctgatggcctggctttgaatgctc

tcatccatagtcataggccagacctatttgactggaatagtgtggtttgccagcagtcagccacacaacgactggaacat

gcattcaacatcgccagatatcaattaggcatagagaaactactcgatcctgaagatgttgataccacctatccagataa

gaagtccatcttaatgtacatcacatcactcttccaagttttgcctcaacaagtgagcattgaagccatccaggaagtgg

aaatgttgccaaggccacctaaagtgactaaagaagaacattttcagttacatcatcaaatgcactattctcaacagatc

acggtcagtctagcacagggatatgagagaacttcttcccctaagcctcgattcaagagctatgcctacacacaggctgc

ttatgtcaccacctctgaccctacacggagcccatttccttcacagcatttggaagctcctgaagacaagtcatttggca

gttcattgatggagagtgaagtaaacctggaccgttatcaaacagctttagaagaagtattatcgtggcttctttctgct

gaggacacattgcaagcacaaggagagatttctaatgatgtggaagtggtgaaagaccagtttcatactcatgaggggta

catgatggatttgacagcccatcagggccgggttggtaatattctacaattgggaagtaagctgattggaacaggaaaat

tatcagaagatgaagaaactgaagtacaagagcagatgaatctcctaaattcaagatgggaatgcctcagggtagctagc

atggaaaaacaaagcaatttacatagagttttaatggatctccagaatcagaaactgaaagagttgaatgactggctaac

aaaaacagaagaaagaacaaggaaaatggaggaagagcctcttggacctgatcttgaagacctaaaacgccaagtacaac

aacataaggtgcttcaagaagatctagaacaagaacaagtcagggtcaattctctcactcacatggtggtggtagttgat

gaatctagtggagatcacgcaactgctgctttggaagaacaacttaaggtattgggagatcgatgggcaaacatctgtag

atggacagaagaccgctgggttcttttacaagacatccttctcaaatggcaacgtcttactgaagaacagtgccttttta

gtgcatggctttcagaaaaagaagatgcagtgaacaagattcacacaactggctttaaagatcaaaatgaaatgttatca

agtcttcaaaaactggccgttttaaaagcggatctagaaaagaaaaagcaatccatgggcaaactgtattcactcaaaca

agatcttctttcaacactgaagaataagtcagtgacccagaagacggaagcatggctggataactttgcccggtgttggg

ataatttagtccaaaaacttgaaaagagtacagcacagactcatagattactgcaacagttccccctggacctggaaaag

ttcttgcctggcttacagaagctgaaacaactgccaatgtcctacaggatgctacccgtaaggaaaggctcctagaaga

ctccaagggagtaaaagagctgatgaaacaatggcaagacctccaaggtgaaattgaagctcacacagatgtttatcaca

acctggatgaaaacagccaaaaaatcctgagatccctggaaggttccgatgatgcagtcctgttacaaagacgtttggat

aacatgaacttcaagtggagtgaacttcggaaaaagtctctcaacattaggtcccatttggaagccagttctgaccagtg

gaagcgtctgcacctttctctgcaggaacttctggtgtggctacagctgaaagatgatgaattaagccggcaggcaccta

ttggaggcgactttccagcagttcagaagcagaacgatgtacatagggccttcaagagggaattgaaaactaaagaacct

gtaatcatgagtactcttgagactgtacgaatatttctgacagagcagcctttggaaggactagagaaactctaccagga

gcccagagagctgcctcctgaggagagagcccagaatgtcactcggcttctacgaaagcaggctgaggaggtcaatactg

agtgggaaaaattgaacctgcactccgctgactggcagagaaaaatagatgagacccttgaaagactccaggaacttcaa

gaggccacggatgagctggacctcaagctgcgccaagctgaggtgatcaagggatcctggcagcccgtgggcgatctcct

cattgactctctccaagatcacctcgagaaagtcaaggcacttcgaggagaaattgcgcctctgaaagagaacgtgagcc

acgtcaatgaccttgctcgccagcttaccactttgggcattcagctctcaccgtataacctcagcactctggaagacctg

aacaccagatggaagcttctgcaggtggccgtcgaggaccgagtcaggcagctgcatgaagcccacagggactttggtcc

agcatctcagcactttctttccacgtctgtccagggtccctgggagagagccatctcgccaaacaaagtgccctactata

tcaaccacgagactcaaacaacttgctgggaccatcccaaaatgacagagctctaccagtctttagctgacctgaataat

gtcagattctcagcttataggactgccatgaaactccgaagactgcagaaggccctttgcttggatctcttgagcctgtc

agctgcatgtgatgccttggaccagcacaacctcaagcaaaatgaccagcccatggatatcctgcagattattaattgtt

tgaccactatttatgaccgcctggagcaagagcacaacaatttggtcaacgtccctctctgcgtggatatgtgtctgaac

tggctgctgaatgtttatgatacgggacgaacagggaggatccgtgtcctgtcttttaaaactggcatcatttccctgtg

taaagcacatttggaagacaagtacagatacctttcaagcaagtggcaagttcaacaggattttgtgaccagcgcaggc

tgggcctccttctgcatgattctatccaaattccaagacagttgggtgaagttgcatcctttgggggcagtaacattgag

ccaagtgtccggagctgcttccaatttgctaataataagccagagatcgaagcggccctcttcctagactggatgagact

ggaaccccagtccatggtgtggctgcccgtcctgcacagagtggctgctgcagaaactgccaagcatcaggccaaatgta

acatctgcaaagagtgtccaatcattggattcaggtacaggagtctaaagcactttaattatgacatctgccaaagctgc

ttttttctggtcgagttgcaaaaggccataaaatgcactatcccatggtggaatattgcactccgactacatcaggaga

agatgttcgagactttgccaaggtactaaaaaacaaatttcgaaccaaaaggtattttgcgaagcatccccgaatgggct

acctgccagtgcagactgtcttagaggggacaacatggaaactcccgacacaatgtagtcgagaggcctaataaagagc

tcagatgcatcgatcagagtgtgttggttttttgtgtgaGATCTAGGAACCCCTAGTGATGGAGTTGGCCACT
CCCTCTC

TGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACC

TTTGGTCGCCCGGCCTCAGTGAGC

GAGCGAGCGCGCAGAGAGGGAGTGGCCAA

Seq. ID No 5: AAV-MCK-Δ3990. It is an AAV vector containing a MCK promoter driving a mini-dystrophin gene connected to a small polyA signal

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTC

GCCCGACGCCCGGGCTTTGCCCG

GGCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTA

GGGGTTCCTAGATCTGAATTCGag

cttgcatgcccactacgggtctaggctgcccatgtaaggaggcaaggcctggggacacccgagatgcctggttataatta

acccagacatgtggctgccccccccccccccaacacctgctgcctgagcctcacccccaccccggtgcctgggtcttaggc

tctgtacaccatggaggagaagctcgctctaaaaataaccctgtccctggtggatcccctgcatgcccaatcaaggctgt

gggggactgagggcaggctgtaacaggcttgggggccagggcttatacgtgcctgggactcccaaagtattactgttcca

tgttcccggcgaagggccagctgtcccccgccagctagactcagcacttagtttaggaaccagtgagcaagtcagcccct

ggggcagcccatacaaggccatggggctgggcaagctgcacgcctgggtccggggtgggcacggtgcccgggcaacgagc

tgaaagctcatctgctctcaggggcccctccctggggacagcccctcctggctagtcacaccctgtaggctcctctatat

aacccaggggcacaggggctgcccccgggtcacTCGAAttttcaCCatgGtttggtgggaagaagtagaggactgttatg

aaagagaagatgttcaaaagaaaacattcacaaaatgggtaaatgcacaattttctaagtttgggaagcagcatattgag

aacctcttcagtgacctacaggatgggaggcgcctcctagacctcctcgaaggcctgacagggcaaaaactgccaaaaga

aaaaggatccacaagagttcatgccctgaacaatgtcaacaaggcactgcgggttttgcagaacaataatgttgatttag

tgaatattggaagtactgacatcgtagatggaaatcataaactgactcttggtttgatttggaatataatcctccactgg

caggtcaaaaatgtaatgaaaaatatcatggctggattgcaacaaaccaacagtgaaaagattctcctgagctgggtccg

acaatcaactcgtaattatccacaggttaatgtaatcaacttcaccaccagctggtctgatggcctggctttgaatgctc

tcatccatagtcataggccagacctatttgactggaatagtgtggtttgccagcagtcagccacacaacgactggaacat

gcattcaacatcgccagatatcaattaggcatagagaaactactcgatcctgaagatgttgataccacctatccagataa

gaagtccatcttaatgtacatcacatcactcttccaagttttgcctcaacaagtgagcattgaagccatccaggaagtgg

aaatgttgccaaggccacctaaagtgactaaagaagaacattttcagttacatcatcaaatgcactattctcaacagatc

acggtcagtctagcacagggatatgagagaacttcttcccctaagcctcgattcaagagctatgcctacacacaggctgc

ttatgtcaccacctctgaccctacacggagcccatttccttcacagcatttggaagctcctgaagacaagtcatttggca

gttcattgatggagagtgaagtaaacctggaccgttatcaaacagctttagaagaagtattatcgtggcttctttctgct

gaggacacattgcaagcacaaggagagatttctaatgatgtggaagtggtgaaagaccagtttcatactcatgaggggta

catgatggatttgacagcccatcagggccgggttggtaatattctacaattgggaagtaagctgattggaacaggaaaat

tatcagaagatgaagaaactgaagtacaagagcagatgaatctcctaaattcaagatgggaatgcctcagggtagctagc

atggaaaaacaaagcaatttacatagagttttaatggatctccagaatcagaaactgaaagagttgaatgactggctaac

aaaaacagaagaaagaacaaggaaaatggaggaagagcctcttggacctgatcttgaagacctaaaacgccaagtacaac

aacataaggtgcttcaagaagatctagaacaagaacaagtcagggtcaattctctcactcacatggtggtggtagttgat

gaatctagtggagatcacgcaactgctgctttggaagaacaacttaaggtattgggagatcgatgggcaaacatctgtag

atggacagaagaccgctgggttcttttacaagaccagcctgacctagctcctggactgaccactattggagcctctccta

ctcagactgttactctggtgacacaacctgtggttactaaggaaactgccatctccaaactagaaatgccatcttccttg

atgttggaggtacctactcatagattactgcaacagttcccccctggacctggaaaagtttcttgcctggcttacagaagc

tgaaacaactgccaatgtcctacaggatgctacccgtaaggaaaggctcctagaagactccaagggagtaaaagagctga

tgaaacaatggcaagacctccaaggtgaaattgaagctcacacagatgtttatcacaacctggatgaaaacagccaaaaa

atcctgagatccctggaaggttccgatgatgcagtcctgttacaaagacgtttggataacatgaacttcaagtggagtga

acttcggaaaaagtctctcaacattaggtcccatttggaagccagttctgaccagtggaagcgtctgcaccttctctgc

aggaacttctggtgtggctacagctgaaagatgatgaattaagccggcaggcacctattggaggcgactttccagcagtt

cagaagcagaacgatgtacatagggccttcaagagggaattgaaaactaaagaacctgtaatcatgagtactcttgagac

tgtacgaatatttctgacagagcagcctttggaaggactagagaaactctaccaggagcccagagagctgcctcctgagg

agagagcccagaatgtcactcggcttctacgaaagcaggctgaggaggtcaatactgagtgggaaaaattgaacctgcac

tccgctgactggcagagaaaaatagatgagacccttgaaagactccaggaacttcaagaggccacggatgagctggacct

caagctgcgccaagctgaggtgatcaagggatcctggcagcccgtgggcgatctcctcattgactctctccaagatcacc

tcgagaaagtcaaggcacttcgaggagaaattgcgcctctgaaagagaacgtgagccacgtcaatgaccttgctcgccag

cttaccactttgggcattcagctctcaccgtataacctcagcactctggaagacctgaacaccagatggaagcttctgca

ggtggccgtcgaggaccgagtcaggcagctgcatgaagcccacagggactttggtccagcatctcagcactttcttcca

cgtctgtccagggtccctgggagagagccatctcgccaaacaaagtgccctactatatcaaccacgagactcaaacaact

tgctgggaccatcccaaaatgacagagctctaccagtctttagctgacctgaataatgtcagattctcagcttataggac

tgccatgaaactccgaagactgcagaaggccctttgcttggatctcttgagcctgtcagctgcatgtgatgccttggacc

agcacaacctcaagcaaaatgaccagcccatggatatcctgcagattattaattgtttgaccactatttatgaccgcctg

gagcaagagcacaacaatttggtcaacgtccctctctgcgtggatatgtgtctgaactggctgctgaatgtttatgatac

gggacgaacagggaggatccgtgtcctgtcttttaaaactggcatcatttccctgtgtaaagcacatttggaagacaagt

acagatacctttttcaagcaagtggcaagttcaacaggattttgtgaccagcgcaggctgggcctccttctgcatgattct

atccaaattccaagacagttgggtgaagttgcatcctttgggggcagtaacattgagccaagtgtccggagctgcttcca

atttgctaataataagccagagatcgaagcggccctcttcctagactggatgagactggaaccccagtccatggtgtggc

tgcccgtcctgcacagagtggctgctgcagaaactgccaagcatcaggccaaatgtaacatctgcaaagagtgtccaatc

attggattcaggtacaggagtctaaagcactttaattatgacatctgccaaagctgctttttttctggtcgagttgcaaa

aggccataaaatgcactatcccatggtggaatattgcactccgactacatcaggagaagatgttcgagacttgccaagg

tactaaaaaaacaaatttcgaaccaaaaggtattttgcgaagcatccccgaatgggctacctgccagtgcagactgtctta

gaggggacaacatggaaactcccgacacaatgtagtcgagaggcctaataaagagctcagatgcatcgatcagagtgtg

ttggtttttgtgtgaGATCTAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCT
CGCTCGCTCACTGAG

GCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAG
CGAGCGAGCGCGCAGAGAGGGAGT

GGCCAA

Seq. ID No 6: AAV-MCK-Δ3849. It is an AAV vector containing a MCK promoter driving a mini-dystrophin gene connected to a small polyA signal

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTC
GCCCGACGCCCGGGCTTTGCCCG

GGCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTA
GGGGTTCCTAGATCTGAATTCGag

cttgcatgcccactacgggtctaggctgcccatgtaaggaggcaaggcctggggacacccgagatgcctggttataatta

acccagacatgtggctgccccccccccccccaacacctgctgcctgagcctcacccccacccggtgcctgggtcttaggc

tctgtacaccatggaggagaagctcgctctaaaaataaccctgtccctggtggatcccctgcatgcccaatcaaggctgt

gggggactgagggcaggctgtaacaggcttgggggccagggcttatacgtgcctgggactcccaaagtattactgttcca

tgttcccggcgaagggccagctgtcccccgccagctagactcagcacttagtttaggaaccagtgagcaagtcagccctt

ggggcagcccatacaaggccatggggctgggcaagctgcacgcctgggtccggggtgggcacggtgcccgggcaacgagc

tgaaagctcatctgctctcaggggcccctccctggggacagcccctcctggctagtcacaccctgtaggctcctctatat

aacccaggggcacaggggctgcccccgggtcacTCGAAttttcaCCatgGtttggtgggaagaagtagaggactgttatg

aaagagaagatgttcaaaagaaaacattcacaaaatgggtaaatgcacaatttctaagtttgggaagcagcatattgag

aacctcttcagtgacctacaggatgggaggcgcctcctagacctcctcgaaggcctgacagggcaaaaactgccaaaaga

aaaaggatccacaagagttcatgccctgaacaatgtcaacaaggcactgcgggttttgcagaacaataatgttgatttag

tgaatattggaagtactgacatcgtagatggaaatcataaactgactcttggtttgatttggaatataatcctccactgg

caggtcaaaaatgtaatgaaaaatatcatggctggattgcaacaaaccaacagtgaaaagattctcctgagctgggtccg

acaatcaactcgtaattatccacaggttaatgtaatcaacttcaccaccagctggtctgatggcctggctttgaatgctc

tcatccatagtcataggccagacctatttgactggaatagtgtggtttgccagcagtcagccacacaacgactggaacat

gcattcaacatcgccagatatcaattaggcatagagaaactactcgatcctgaagatgttgataccacctatccagataa

gaagtccatcttaatgtacatcacatcactcttccaagttttgcctcaacaagtgagcattgaagccatccaggaagtgg

aaatgttgccaaggccacctaaagtgactaaagaagaacattttcagttacatcatcaaatgcactattctcaacagatc

acggtcagtctagcacagggatatgagagaacttcttcccctaagcctcgattcaagagctatgcctacacacaggctgc

ttatgtcaccacctctgaccctacacggagcccatttccttcacagcatttggaagctcctgaagacaagtcatttggca

gttcattgatggagagtgaagtaaacctggaccgttatcaaacagctttagaagaagtattatcgtggcttctttctgct

gaggacacattgcaagcacaaggagagatttctaatgatgtggaagtggtgaaagaccagtttcatactcatgaggggta

catgatggatttgacagcccatcagggccgggttggtaatattctacaattgggaagtaagctgattggaacaggaaaat

tatcagaagatgaagaaactgaagtacaagagcagatgaatctcctaaattcaagatgggaatgcctcagggtagctagc

atggaaaaacaaagcaatttacatagagttttaatggatctccagaatcagaaactgaaagagttgaatgactggctaac

aaaaacagaagaaagaacaaggaaaatggaggaagagcctcttggacctgatcttgaagacctaaaacgccaagtacaac

aacataaggtgcttcaagaagatctagaacaagaacaagtcagggtcaattctctcactcacatggtggtggtagttgat

gaatctagtggagatcacgcaactgctgctttggaagaacaacttaaggtattgggagatcgatgggcaaacatctgtag

atggacagaagaccgctgggttcttttacaagacactcatagattactgcaacagttccccctggacctggaaaagtttc

ttgcctggcttacagaagctgaaacaactgccaatgtcctacaggatgctacccgtaaggaaaggctcctagaagactcc

aagggagtaaaagagctgatgaaacaatggcaagacctccaaggtgaaattgaagctcacacagatgtttatcacaacct

ggatgaaaacagccaaaaaatcctgagatccctggaaggttccgatgatgcagtcctgttacaaagacgtttggataaca

tgaacttcaagtggagtgaacttcggaaaaagtctctcaacattaggtcccatttggaagccagttctgaccagtggaag

cgtctgcacctttctctgcaggaacttctggtgtggctacagctgaaagatgatgaattaagccggcaggcacctattgg

aggcgactttccagcagttcagaagcagaacgatgtacatagggccttcaagagggaattgaaaactaaagaacctgtaa

tcatgagtactcttgagactgtacgaatatttctgacagagcagcctttggaaggactagagaaactctaccaggagccc

agagagctgcctcctgaggagagagcccagaatgtcactcggcttctacgaaagcaggctgaggaggtcaatactgagtg

ggaaaaattgaacctgcactccgctgactggcagagaaaaatagatgagacccttgaaagactccaggaacttcaagagg

ccacggatgagctggacctcaagctgcgccaagctgaggtgatcaagggatcctggcagcccgtgggcgatctcctcatt

gactctctccaagatcacctcgagaaagtcaaggcacttcgaggagaaattgcgcctctgaaagagaacgtgagccacgt

caatgaccttgctcgccagcttaccactttgggcattcagctctcaccgtataacctcagcactctggaagacctgaaca

ccagatggaagcttctgcaggtggccgtcgaggaccgagtcaggcagctgcatgaagcccacagggactttggtccagca

tctcagcactttctttccacgtctgtccagggtccctgggagagagccatctcgccaaacaaagtgccctactatatcaa

ccacgagactcaaacaacttgctgggaccatcccaaaatgacagagctctaccagtctttagctgacctgaataatgtca

gattctcagcttataggactgccatgaaactccgaagactgcagaaggccctttgcttggatctcttgagcctgtcagct

gcatgtgatgccttggaccagcacaacctcaagcaaaatgaccagcccatggatatcctgcagattattaattgtttgac

cactatttatgaccgcctggagcaagagcacaacaatttggtcaacgtccctctctgcgtggatatgtgtctgaactggc

tgctgaatgtttatgatacgggacgaacagggaggatccgtgtcctgtcttttaaaactggcatcatttccctgtgtaaa

gcacatttggaagacaagtacagataccttttcaagcaagtggcaagttcaacaggattttgtgaccagcgcaggctggg

cctccttctgcatgattctatccaaattccaagacagttgggtgaagttgcatcctttgggggcagtaacattgagccaa

gtgtccggagctgcttccaatttgctaataataagccagagatcgaagcggccctcttcctagactggatgagactggaa

ccccagtccatggtgtggctgcccgtcctgcacagagtggctgctgcagaaactgccaagcatcaggccaaatgtaacat

ctgcaaagagtgtccaatcattggattcaggtacaggagtctaaagcactttaattatgacatctgccaaagctgctttt

ttctggtcgagttgcaaaaggccataaaatgcactatcccatggtggaatattgcactccgactacatcaggagaagat

gttcgagactttgccaaggtactaaaaaacaaatttcgaaccaaaaggtattttgcgaagcatccccgaatgggctacct

gccagtgcagactgtcttagaggggacaacatggaaactcccgacacaatgtagtcgagaggcctaataaagagctcag

atgcatcgatcagagtgtgttggtttttgtgtgaGATCTAGGAACCCCTAGTGATGGAGTTGGCCACTCCC
TCTCTGCG

CGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTG
GTCGCCCGGCCTCAGTGAGCGAGC

GAGCGCGCAGAGAGGGAGTGGCCAA

Seq. ID No 7: AAV-CMV-Δ3990. It is an AAV vector containing a CMV promoter driving a mini-dystrophin gene connected to a small polyA signal

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTC
GCCCGACGCCCGGGCTTTGCCCG

GGCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTA
GGGGTTCCTAGATCTGAATTCGGT

ACCCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCG
CCCATTGACGTCAATAATGACGT

ATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATT
TACGGTAAACTGCCCACTTGGCA

GTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAA
TGGCCCGCCTGGCATTATGCCCA

GTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCT
ATTACCATGGTGATGCGGTTTT

GGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCC
ACCCCATTGACGTCAATGGGAGT

TTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCA
TTGACGCAAATGGGCGGTAGGCG

TGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTG
GAGACGCCATCCACGCTGTTTTG

ACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCGGACTCTAGAGGATCCGGTA
CtcgaAttttcaCCatgGtttggt

gggaagaagtagaggactgttatgaaagagaagatgttcaaaagaaaacattcacaaaatgggtaaatgcacaattttct

aagtttgggaagcagcatattgagaacctcttcagtgacctacaggatgggaggcgcctcctagacctcctcgaaggcct

gacagggcaaaaactgccaaaagaaaaaggatccacaagagttcatgccctgaacaatgtcaacaaggcactgcgggttt

tgcagaacaataatgttgatttagtgaatattggaagtactgacatcgtagatggaaatcataaactgactcttggtttg

atttggaatataatcctccactggcaggtcaaaaatgtaatgaaaaatatcatggctggattgcaacaaaccaacagtga

aaagattctcctgagctgggtccgacaatcaactcgtaattatccacaggttaatgtaatcaacttcaccaccagctggt

ctgatggcctggctttgaatgctctcatccatagtcataggccagacctatttgactggaatagtgtggtttgccagcag

tcagccacacaacgactggaacatgcattcaacatcgccagatatcaattaggcatagagaaactactcgatcctgaaga

tgttgataccacctatccagataagaagtccatcttaatgtacatcacatcactcttccaagttttgcctcaacaagtga

gcattgaagccatccaggaagtggaaatgttgccaaggccacctaaagtgactaaagaagaacattttcagttacatcat

caaatgcactattctcaacagatcacggtcagtctagcacagggatatgagagaacttcttcccctaagcctcgattcaa

gagctatgcctacacacaggctgcttatgtcaccacctctgaccctacacggagcccatttccttcacagcatttggaag

ctcctgaagacaagtcatttggcagttcattgatggagagtgaagtaaacctggaccgttatcaaacagctttagaagaa

gtattatcgtggcttctttctgctgaggacacattgcaagcacaaggagagatttctaatgatgtggaagtggtgaaaga

ccagtttcatactcatgaggggtacatgatggatttgacagcccatcagggccgggttggtaatattctacaattgggaa

gtaagctgattggaacaggaaaattatcagaagatgaagaaactgaagtacaagagcagatgaatctcctaaattcaaga

tgggaatgcctcagggtagctagcatggaaaaacaaagcaatttacatagagttttaatggatctccagaatcagaaact

gaaagagttgaatgactggctaacaaaaacagaagaaagaacaaggaaaatggaggaagagcctcttggacctgatcttg

aagacctaaaacgccaagtacaacaacataaggtgcttcaagaagatctagaacaagaacaagtcagggtcaattctctc

actcacatggtggtggtagttgatgaatctagtggagatcacgcaactgctgctttggaagaacaacttaaggtattggg

agatcgatgggcaaacatctgtagatggacagaagaccgctgggttcttttacaagaccagcctgacctagctcctggac

tgaccactattggagcctctcctactcagactgttactctggtgacacaacctgtggttactaaggaaactgccatctcc

aaactagaaatgccatcttccttgatgttggaggtacctactcatagattactgcaacagttccccctggacctggaaaa

gtttcttgcctggcttacagaagctgaaacaactgccaatgtcctacaggatgctacccgtaaggaaaggctcctagaag

actccaagggagtaaaagagctgatgaaacaatggcaagacctccaaggtgaaattgaagctcacacagatgtttatcac

aacctggatgaaaacagccaaaaaatcctgagatccctggaaggttccgatgatgcagtcctgttacaaagacgtttgga

taacatgaacttcaagtggagtgaacttcggaaaaagtctctcaacattaggtcccatttggaagccagttctgaccagt

ggaagcgtctgcacctttctctgcaggaacttctggtgtggctacagctgaaagatgatgaattaagccggcaggcacct

attggaggcgactttccagcagttcagaagcagaacgatgtacatagggccttcaagagggaattgaaaactaaagaacc

tgtaatcatgagtactcttgagactgtacgaatatttctgacagagcagcctttggaaggactagagaaactctaccagg

agcccagagagctgcctcctgaggagagagcccagaatgtcactcggcttctacgaaagcaggctgaggaggtcaatact

gagtgggaaaaattgaacctgcactccgctgactggcagagaaaaatagatgagacccttgaaagactccaggaacttca

agaggccacggatgagctggacctcaagctgcgccaagctgaggtgatcaagggatcctggcagcccgtgggcgatctcc

tcattgactctctccaagatcacctcgagaaagtcaaggcacttcgaggagaaattgcgcctctgaaagagaacgtgagc

cacgtcaatgaccttgctcgccagcttaccactttgggcattcagctctcaccgtataacctcagcactctggaagacct

gaacaccagatggaagcttctgcaggtggccgtcgaggaccgagtcaggcagctgcatgaagcccacagggactttggtc

cagcatctcagcactttctttccacgtctgtccagggtccctgggagagagccatctcgccaaacaaagtgccctactat

atcaaccacgagactcaaacaacttgctgggaccatcccaaaatgacagagctctaccagtctttagctgacctgaataa

tgtcagattctcagcttataggactgccatgaaactccgaagactgcagaaggccctttgcttggatctcttgagcctgt

cagctgcatgtgatgccttggaccagcacaacctcaagcaaaatgaccagcccatggatatcctgcagattattaattgt

ttgaccactatttatgaccgcctggagcaagagcacaacaatttggtcaacgtccctctctgcgtggatatgtgtctgaa

ctggctgctgaatgtttatgatacgggacgaacagggaggatccgtgtcctgtcttttaaaactggcatcatttccctgt

gtaaagcacatttggaagacaagtacagataccttttcaagcaagtggcaagttcaacaggattttgtgaccagcgcagg

ctgggcctccttctgcatgattctatccaaattccaagacagttgggtgaagttgcatcctttgggggcagtaacattga

gccaagtgtccggagctgcttccaatttgctaataataagccagagatcgaagcggccctcttcctagactggatgagac

tggaaccccagtccatggtgtggctgcccgtcctgcacagagtggctgctgcagaaactgccaagcatcaggccaaatgt

aacatctgcaaagagtgtccaatcattggattcaggtacaggagtctaaagcactttaattatgacatctgccaaagctg

ctttttttctggtcgagttgcaaaaggccataaaatgcactatcccatggtggaatattgcactccgactacatcaggag

aagatgttcgagactttgccaaggtactaaaaaacaaatttcgaaccaaaaggtattttgcgaagcatccccgaatgggc

tacctgccagtgcagactgtcttagaggggacaacatggaaactcccgacacaatgtagtcgagaggcctaataaagag

ctcagatgcatcgatcagagtgtgttggtttttgtgtgaGATCTAGGAACCCCTAGTGATGGAGTTGGCCACT
CCCTCT

CTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGAC
CTTTGGTCGCCCGGCCTCAGTGAG

CGAGCGAGCGCGCAGAGAGGGAGTGGCCAA

Seq. ID No 8: AAV-CMV-Δ3849. It is an AAV vector containing a CMV promoter driving a mini-dystrophin gene connected to a small polyA signal

TGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCG
CCCGACGCCCGGGCTTTGCCCGG

GCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAG
GGGTTCCTAGATCTGAATTCGGTA

CCCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGC
CCATTGACGTCAATAATGACGTA

TGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTT
ACGGTAAACTGCCCACTTGGCAG

TACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAAT
GGCCCGCCTGGCATTATGCCCAG

TACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTA
TTACCATGGTGATGCGGTTTTG

GCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCA
CCCCATTGACGTCAATGGGAGTT

TGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCAT
TGACGCAAATGGGCGGTAGGCGT

GTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGG
AGACGCCATCCACGCTGTTTTGA

CCTCCATAGAAGACACCGGGACCGATCCAGCCTCCGGACTCTAGAGGATCCGGTACt
cgaAttttcaCCatgGttggtg

ggaagaagtagaggactgttatgaaagagaagatgttcaaaagaaaacattcacaaaatgggtaaatgcacaattttcta

agtttgggaagcagcatattgagaacctcttcagtgacctacaggatgggaggcgcctcctagacctcctcgaaggcctg

acagggcaaaaactgccaaaagaaaaaggatccacaagagttcatgccctgaacaatgtcaacaaggcactgcgggtttt

gcagaacaataatgttgatttagtgaatattggaagtactgacatcgtagatggaaatcataaactgactcttggtttga

tttggaatataatcctccactggcaggtcaaaaatgtaatgaaaaatatcatggctggattgcaacaaaccaacagtgaa

aagattctcctgagctgggtccgacaatcaactcgtaattatccacaggttaatgtaatcaacttcaccaccagctggtc

tgatggcctggctttgaatgctctcatccatagtcataggccagacctatttgactggaatagtgtggtttgccagcagt

cagccacacaacgactggaacatgcattcaacatcgccagatatcaattaggcatagagaaactactcgatcctgaagat

gttgataccacctatccagataagaagtccatcttaatgtacatcacatcactcttccaagttttgcctcaacaagtgag

cattgaagccatccaggaagtggaaatgttgccaaggccacctaaagtgactaaagaagaacattttcagttacatcatc

aaatgcactattctcaacagatcacggtcagtctagcacagggatatgagagaacttcttcccctaagcctcgattcaag

agctatgcctacacacaggctgcttatgtcaccacctctgaccctacacggagcccatttccttcacagcatttggaagc

tcctgaagacaagtcatttggcagttcattgatggagagtgaagtaaacctggaccgttatcaaacagctttagaagaag

tattatcgtggcttctttctgctgaggacacattgcaagcacaaggagagatttctaatgatgtggaagtggtgaaagac

cagtttcatactcatgaggggtacatgatggatttgacagcccatcagggccgggttggtaatattctacaattgggaag

taagctgattggaacaggaaaattatcagaagatgaagaaactgaagtacaagagcagatgaatctcctaaattcaagat

gggaatgcctcagggtagctagcatggaaaaacaaagcaatttacatagagttttaatggatctccagaatcagaaactg

aaagagttgaatgactggctaacaaaaacagaagaaagaacaaggaaaatggaggaagagcctcttggacctgatcttga

agacctaaaacgccaagtacaacaacataaggtgcttcaagaagatctagaacaagaacaagtcagggtcaattctctca

ctcacatggtggtggtagttgatgaatctagtggagatcacgcaactgctgctttggaagaacaacttaaggtattggga

gatcgatgggcaaacatctgtagatggacagaagaccgctgggttcttttacaagacactcatagattactgcaacagtt

ccccctggacctggaaaagtttcttgcctggcttacagaagctgaaacaactgccaatgtcctacaggatgctacccgta

aggaaaggctcctagaagactccaagggagtaaaagagctgatgaaacaatggcaagacctccaaggtgaaattgaagct

cacacagatgtttatcacaacctggatgaaaacagccaaaaaatcctgagatccctggaaggttccgatgatgcagtcct

gttacaaagacgtttggataacatgaacttcaagtggagtgaacttcggaaaaagtctctcaacattaggtcccatttgg

aagccagttctgaccagtggaagcgtctgcaccttctctgcaggaacttctggtgtggctacagctgaaagatgatgaa

ttaagccggcaggcacctattggaggcgactttccagcagttcagaagcagaacgatgtacatagggccttcaagaggga

attgaaaactaaagaacctgtaatcatgagtactcttgagactgtacgaatatttctgacagagcagcctttggaaggac

tagagaaactctaccaggagcccagagagctgcctcctgaggagagagcccagaatgtcactcggcttctacgaaagcag

gctgaggaggtcaatactgagtgggaaaaattgaacctgcactccgctgactggcagagaaaaatagatgagacccttga

aagactccaggaacttcaagaggccacggatgagctggacctcaagctgcgccaagctgaggtgatcaagggatcctggc

agcccgtgggcgatctcctcattgactctctccaagatcacctcgagaaagtcaaggcacttcgaggagaaattgcgcct

ctgaaagagaacgtgagccacgtcaatgaccttgctcgccagcttaccactttgggcattcagctctcaccgtataacct

cagcactctggaagacctgaacaccagatggaagcttctgcaggtggccgtcgaggaccgagtcaggcagctgcatgaag

cccacagggactttggtccagcatctcagcactttctttccacgtctgtccagggtccctgggagagagccatctcgcca

aacaaagtgccctactatatcaaccacgagactcaaacaacttgctgggaccatcccaaaatgacagagctctaccagtc

tttagctgacctgaataatgtcagattctcagcttataggactgccatgaaactccgaagactgcagaaggccctttgct

tggatctcttgagcctgtcagctgcatgtgatgccttggaccagcacaacctcaagcaaaatgaccagcccatggatatc

ctgcagattattaattgtttgaccactatttatgaccgcctggagcaagagcacaacaatttggtcaacgtccctctctg

cgtggatatgtgtctgaactggctgctgaatgtttatgatacgggacgaacagggaggatccgtgtcctgtctttaaaa

ctggcatcatttccctgtgtaaagcacatttggaagacaagtacagatacctttcaagcaagtggcaagttcaacagga

ttttgtgaccagcgcaggctgggcctccttctgcatgattctatccaaattccaagacagttgggtgaagttgcatcctt

tgggggcagtaacattgagccaagtgtccggagctgcttccaatttgctaataataagccagagatcgaagcggccctct

tcctagactggatgagactggaaccccagtccatggtgtggctgcccgtcctgcacagagtggctgctgcagaaactgcc

aagcatcaggccaaatgtaacatctgcaaagagtgtccaatcattggattcaggtacaggagtctaaagcactttaatta

tgacatctgccaaagctgctttttttctggtcgagttgcaaaaggccataaaatgcactatcccatggtggaatattgca

ctccgactacatcaggagaagatgttcgagactttgccaaggtactaaaaaacaaatttcgaaccaaaaggtattttgcg

aagcatccccgaatgggctacctgccagtgcagactgtcttagaggggacaacatggaaactcccgacacaatgtagtc

gagaggcctaataaagagctcagatgcatcgatcagagtgtgttggtttttgtgtgaGATCTAGGAACCCCTAGTGATG

GAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAG
CCCGGGCGTCGGGCGACCTTTGG

TCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAA

Seq. ID No 9: AAV-E-CMV-Δ3849: It is an AAV vector containing a MCK enhancer with a CMV promoter driving a mini-dystrophin gene connected to a small polyA signal

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTC
GCCCGACGCCCGGGCTTTGCCCG

GGCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTA
GGGGTTCCTAGATCTGAATTCGGT

ACcactacgggtctaggctgcccatgtaaggaggcaaggcctggggacacccgagatgcctggttataattaacccagac

atgtggctgccccccccccccccaacacctgctgcctgagcctcacccccaccccggtgcctgggtcttaggctctgtaca

ccatggaggagaagctcgctctaaaaataaccctgtccctggtggatcGGTACCCGTTACATAACTTACGGTAAA
TGGCC

CGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTC
CCATAGTAACGCCAATAGGGACT

TTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACAT
CAAGTGTATCATATGCCAAGTAC

GCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACAT
GACCTTATGGGACTTTCCTACTT

GGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGT
ACATCAATGGGCGTGGATAGCGG

TTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTT
GGCACCAAAATCAACGGGACTT

TCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTAC
GGTGGGAGGTCTATATAAGCAGAG

CTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTGACCTCC
ATAGAAGACACCGGGACCGATCC

AGCCTCCGGACTCTAGAGGATCCGGTACtcgaAttttcaCCatgGtttggtgggaagaagtagaggactgttat
gaaaga

gaagatgttcaaaagaaaacattcacaaaatgggtaaatgcacaattttctaagtttgggaagcagcatattgagaacct

cttcagtgacctacaggatgggaggcgcctcctagacctcctcgaaggcctgacagggcaaaaactgccaaaagaaaaag

gatccacaagagttcatgccctgaacaatgtcaacaaggcactgcgggttttgcagaacaataatgttgatttagtgaat

attggaagtactgacatcgtagatggaaatcataaactgactcttggtttgatttggaatataatcctccactggcaggt

caaaaatgtaatgaaaaatatcatggctggattgcaacaaaccaacagtgaaaagattctcctgagctgggtccgacaat

caactcgtaattatccacaggttaatgtaatcaacttcaccaccagctggtctgatggcctggctttgaatgctctcatc

catagtcataggccagacctatttgactggaatagtgtggtttgccagcagtcagccacacaacgactggaacatgcatt

caacatcgccagatatcaattaggcatagagaaaactactcgatcctgaagatgttgataccacctatccagataagaagt

ccatcttaatgtacatcacatcactcttccaagttttgcctcaacaagtgagcattgaagccatccaggaagtggaaatg

ttgccaaggccacctaaagtgactaaagaagaacattttcagttacatcatcaaatgcactattctcaacagatcacggt

cagtctagcacagggatatgagagaacttcttcccctaagcctcgattcaagagctatgcctacacacaggctgcttatg

tcaccacctctgaccctacacggagcccatttccttcacagcatttggaagctcctgaagacaagtcatttggcagttca

ttgatggagagtgaagtaaacctggaccgttatcaaacagctttagaagaagtattatcgtggcttctttctgctgagga

cacattgcaagcacaaggagagatttctaatgatgtggaagtggtgaaagaccagtttcatactcatgaggggtacatga

tggatttgacagcccatcagggccgggttggtaatattctacaattgggaagtaagctgattggaacaggaaaattatca

gaagatgaagaaactgaagtacaagagcagatgaatctcctaaattcaagatgggaatgcctcagggtagctagcatgga

aaaacaaagcaatttacatagagttttaatggatctccagaatcagaaactgaaagagttgaatgactggctaacaaaaa

cagaagaaagaacaaggaaaatggaggaagagcctcttggacctgatcttgaagacctaaaacgccaagtacaacaacat

aaggtgcttcaagaagatctagaacaagaacaagtcagggtcaattctctcactcacatggtggtggtagttgatgaatc

tagtggagatcacgcaactgctgctttggaagaacaacttaaggtattgggagatcgatgggcaaacatctgtagatgga

cagaagaccgctgggttcttttacaagacactcatagattactgcaacagttccccctggacctggaaaagtttcttgcc

tggcttacagaagctgaaacaactgccaatgtcctacaggatgctacccgtaaggaaaggctcctagaagactccaaggg

agtaaaagagctgatgaaacaatggcaagacctccaaggtgaaattgaagctcacacagatgtttatcacaacctggatg

aaaacagccaaaaaatcctgagatccctggaaggttccgatgatgcagtcctgttacaaagacgtttggataacatgaac

ttcaagtggagtgaacttcggaaaaagtctctcaacattaggtcccatttggaagccagttctgaccagtggaagcgtct

gcacctttctctgcaggaacttctggtgtggctacagctgaaagatgatgaattaagccggcaggcacctattggaggcg

actttccagcagttcagaagcagaacgatgtacatagggccttcaagagggaattgaaaactaaagaacctgtaatcatg

agtactcttgagactgtacgaatatttctgacagagcagcctttggaaggactagagaaactctaccaggagcccagaga

gctgcctcctgaggagagagcccagaatgtcactcggcttctacgaaagcaggctgaggaggtcaatactgagtgggaaa

aattgaacctgcactccgctgactggcagagaaaaatagatgagacccttgaaagactccaggaacttcaagaggccacg

gatgagctggacctcaagctgcgccaagctgaggtgatcaagggatcctggcagcccgtgggcgatctcctcattgactc

tctccaagatcacctcgagaaagtcaaggcacttcgaggagaaattgcgcctctgaaagagaacgtgagccacgtcaatg

accttgctcgccagcttaccactttgggcattcagctctcaccgtataacctcagcactctggaagacctgaacaccaga

tggaagcttctgcaggtggccgtcgaggaccgagtcaggcagctgcatgaagcccacagggactttggtccagcatctca

gcactttctttccacgtctgtccagggtccctgggagagagccatctcgccaaacaaagtgccctactatatcaaccacg

agactcaaacaacttgctgggaccatcccaaaatgacagagctctaccagtctttagctgacctgaataatgtcagattc

tcagcttataggactgccatgaaactccgaagactgcagaaggccctttgcttggatctcttgagcctgtcagctgcatg

tgatgccttggaccagcacaacctcaagcaaaatgaccagcccatggatatcctgcagattattaattgtttgaccacta

tttatgaccgcctggagcaagagcacaacaatttggtcaacgtccctctctgcgtggatatgtgtctgaactggctgctg

aatgtttatgatacgggacgaacagggaggatccgtgtcctgtcttttaaaactggcatcatttccctgtgtaaagcaca

tttggaagacaagtacagatacctttcaagcaagtggcaagttcaacaggattttgtgaccagcgcaggctgggcctcc

ttctgcatgattctatccaaattccaagacagttgggtgaagttgcatcctttgggggcagtaacattgagccaagtgtc

cggagctgcttccaatttgctaataataagccagagatcgaagcggccctcttcctagactggatgagactggaacccca

gtccatggtgtggctgcccgtcctgcacagagtggctgctgcagaaactgccaagcatcaggccaaatgtaacatctgca

aagagtgtccaatcattggattcaggtacaggagtctaaagcactttaattatgacatctgccaaagctgctttttttct

ggtcgagttgcaaaaggccataaaatgcactatcccatggtggaatattgcactccgactacatcaggagaagatgttcg

agactttgccaaggtactaaaaaacaaatttcgaaccaaaaggtattttgcgaagcatccccgaatgggctacctgccag

tgcagactgtcttagaggggacaacatggaaactcccgacacaatgtagtcgagaggcctaataaagagctcagatgca

tcgatcagagtgtgttggttttttgtgtgaGATCTAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCT
CTGCGCGCTC

GCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGC
CCGGCCTCAGTGAGCGAGCGAGCG

CGCAGAGAGGGAGTGGCCAA

DETAILED DESCRIPTION AND PREFERRED EMBODIMENTS OF THE INVENTION

To explore the feasibility of using AAV vectors for DMD gene therapy, we have devised strategies to create novel truncated dystrophin genes, which are small enough to be packaged into AAV vectors, and yet retain the essential functions needed to protect muscle from the pathological symptoms.

A. Definitions

In describing the present invention, the following terms will be employed, and are intended to be defined as indicated below.

"Gene transfer" or "gene delivery" refers to methods or systems for reliably inserting a particular nucleotide sequence (e.g., DNA) into targeted cells. Such methods preferably result in the integration of the transferred genetic material into the genome of target cells. Gene transfer provides a unique approach for the treatment of acquired and inherited diseases, and a number of systems have been developed in the art for gene transfer into mammalian cells. See, e.g., U.S. Pat. No. 5,399,346.

By "vector" is meant any genetic element, such as a plasmid, phage, transposon, cosmid, chromosome, virus, virion, etc., which is capable of replication when associated with the proper control elements and which can transfer gene sequences between cells. Thus, the term includes cloning and expression vehicles, as well as viral vectors.

By an "AAV vector" is meant a vector derived from an adeno-associated virus serotype, including without limitation, AAV-1, AAV-2, AAV-3, AAV-4, AAV-5, AAVX7, etc. AAV

vectors can have one or more of the AAV wild-type genes deleted in whole or part, preferably the rep and/or cap genes, but retain functional flanking ITR sequences. AAV vectors can be constructed using recombinant techniques that are known in the art to include one or more heterologous nucleotide sequences flanked on both ends (5' and 3') with functional AAV ITRs. In the practice of the invention, an AAV vector can include at least one AAV ITR and a suitable promoter sequence positioned upstream of the heterologous nucleotide sequence and at least one AAV ITR positioned downstream of the heterologous sequence. A "recombinant AAV vector plasmid" refers to one type of recombinant AAV ector wherein the vector comprises a plasmid. As with AAV vectors in general, 5' and 3' ITRs flank the selected heterologous nucleotide sequence. AAV vectors can also include transcription sequences such as polyadenylation sites, as well as selectable markers or reporter genes, enhancer sequences, and other control elements which allow for the induction of transcription. Such control elements are described more fully below.

As used herein, the term "AAV virion" or "AAV particle" refers to a complete virus particle. An AAV virion may be a wild type AAV virus particle (comprising a linear, single-stranded AAV nucleic acid genome associated with an AAV capsid, i.e., a protein coat), or a recombinant AAV virus particle (described below). In this regard, single-stranded AAV nucleic acid molecules (either the sense/coding strand or the antisense/anticoding strand as those terms are generally defined) can be packaged into an AAV virion; both the sense and the antisense strands are equally infectious.

As used herein, the term "recombinant AAV virion", "recombinant AAV particle" or "rAAV" is defined as an infectious, replication-defective virus composed of an AAV protein shell encapsidating (i.e., surrounding with a protein coat) a heterologous nucleotide sequence, which in turn is flanked 5' and 3' by AAV ITRs. A number of techniques for constructing

recombinant AAV virions are known in the art. See, e.g., U.S. Pat. No. 5,173,414; International Publication Numbers WO 92/01070 (published 23 Jan. 1992) and WO 93/03769 (published 4 Mar. 1993); Lebkowski et al. (1988) Molec. Cell. Biol. 8:3988-3996; Vincent et al. (1990) Vaccines 90 (Cold Spring Harbor Laboratory Press); Carter, B.J. (1992) Current Opinion in Biotechnology 3:533-539; Muzyczka, N. (1992) Current Topics in Microbiol. and Immunol. 158:97-129; Kotin, R. M. (1994) Human Gene Therapy 5:793-801; Shelling and Smith (1994) Gene Therapy 1:165-169; and Zhou et al. (1994) J. Exp. Med. 179:1867-1875.

The term "mini-dystrophin gene" is used to refer to the novel dystrophin constructs created by extensive deletions in the central rod domain plus extensive deletion in the C-terminal domain of the human dystrophin cDNA. In addition, the mini-dystrophin genes also contain artificial DNA sequences surrounding the original protein translation initiation codon ATG. The artificial sequences enhances the mini-dystrophin protein synthesis. The mini-dystrophin genes are smaller than the 5-kilobase packaging limit of AAV viral vectors. And most importantly, the mini-dystrophin genes harbor biological functions that can protect the muscle from dystrophic pathology and symptoms.

The symbol "Δ" (delta) is a prefix for the mini-dystrophin genes that contain deletions as described above.

B. General Methods

Dystrophin is an enormous rod-like protein(427 kDa) localized beneath the inner surface of muscle cell membrane. [Watkins, S. C. et al., *Nature* 333, 863-866 (1988)]. It functions through four major structural domains. The N-terminal domain binds to the F-actin of cytoskeletal structures, while the cysteine-rich (CR) domain along with the distal C-terminal

domain anchors to the cell membrane via dystrophin-associated protein (DAP) complexes, thus, dystrophin crosslinks and stabilizes the muscle cell membrane and cytoskeleton. The central rod domain contains 24 triple-helix rod repeats and 4 hinges. [Koenig, M. et al., *J Biol Chem* 265, 4560-4566 (1990)]. It presumably functions as a "shock absorber" during muscle contraction. Interestingly, a number of mild muscular dystrophy patients, although they endure large in-frame deletions in the central rod domain, suffer only slight symptoms. [England, S. B. *et al. Nature* 343, 180-182 (1990); Passos-Bueno, M. R. et al., *Hum Mol Genet* 3, 919-922 (1994); and Mirabella, M. *et al. Neurology* 51, 592-595 (1998)]. This phenomenon suggests that a major portion of the rod domain is dispensable. In addition, transgenic studies in *mdx* mice showed that two deletion mutants in C-terminus, one lacking exons 71-74 and the other lacking exons 75-78, displayed full functions in preventing dystrophic phenotypes. [Rafael, J. A. *et al. Journal of Cell Biology* 134, 93-102 (1996)]. These findings suggest that the distal C-terminal region (exons 71-74 and 75-78) may also be dispensable. In contrast, N-terminal deletions variably impair dystrophin functions. [Corrado, K. *et al. J Cell Biol* 134, 873-884 (1996)]. We have created by rational design several mini-genes, in each deleting up to 3/4 of the central rod domain (19 rods and 2 hinges) and nearly the entire distal C-terminal domain (exons 71-78) (Fig. 1). These mini-genes have enabled us to re-examine a previous hypothesis that a dystrophin could not be made smaller than 50% of the full length without causing muscular dystrophies. [Fanin, M. *et al. Muscle Nerve* 19, 1154-1160 (1996)]. These novel mini-dystrophin genes, representing only one third (1/3) of the 11 kb full-length dystrophin coding sequence, are significantly smaller than the 6.3 kb Becker-form mini-dystrophin [England, S. B. *et al. Nature* 343, 180-182 (1990)] that was previously widely used in transgenic studies and gene therapy studies. [Ragot, T. *et al. Nature* 361, 647-650 (1993); Cox, G. A. *et al. Nature* 364, 725-729 (1993); and Wells, D. J. *et al. Hum Mol Genet* 4, 1245-1250 (1995)]. To ensure sufficient physical flexibility of the protein, all of our mini-dystrophins still retain at least five rod repeats (R1, R2, R22, R23 & R24) and 2 hinges (H1 & H4) in the central rod domain (Fig. 1). Construct Δ4173 has an additional rod (R3), while Δ3990 has an additional hinge (H3) (Fig. 1).

To investigate the functionality of the novel mini-dystrophin constructs, it is essential to demonstrate that they can protect muscle from the dystrophic phenotype. The onset of the phenotype in *mdx* mice starts at around three weeks of age with massive waves of myofiber

degeneration/regeneration. Regenerated myofibers are characterized by the presence of central nuclei, while the myonuclei in normal myofibers are peripherally located. The presence of central nuclei is a primary pathological sign of muscular dystrophies. The absence of central nucleation after gene therapy would suggest that the therapy was successful. However even in healthy mice, a majority of the myonuclei remain centrally located after experiencing a transient pathology such as myotoxin treatment. [Martin, H. et al., *Muscle Nerve* 11, 588-596 (1988)]. Because adult *mdx* muscles already have extensive central nucleation, this makes it complex to evaluate the benefits of gene therapy by judging the status of myonuclei. Therefore, we chose to test the AAV mini-dystrophin constructs in young *mdx* mice (10-day old) before the onset of central nucleation, so that the therapeutic effects of the mini-genes can be evaluated with certainty. Furthermore, equally crucial is the alleviation of other pathological signs, including wide variation of myofiber diameters, round (non-polygonal) myofiber shapes in transverse sections, proliferation of connective tissues (fibrosis), and finally loss of muscle cell membrane integrity.

To investigate the therapeutic effects in *mdx* mice, we injected into the hindleg muscle (gastrocnemius) of 10-day old pups with the novel mini-dystrophin genes, which were packaged into AAV vectors containing an MCK (muscle-specific creatine kinase) promoter [Shield, M. A. et al., *Mol Cell Biol* 16, 5058-5068 (1996)] to ensure muscle-specific expression. Three months after vector injection, the muscles were collected to evaluate mini-dystrophin expression, as well as for biochemical restoration of the DAP complexes, which are absence due to the primary deficiency of dystrophin. Immunofluorescent staining on thin sections of AAV treated muscles, using an antibody (Dys3) specific to the human dystrophin N-terminal region, revealed wide-spread vector transduction and correct submembrane location of the mini-dystrophins in a majority of the myofibers, especially in muscles treated with vector AAV-MCK-3849 or AAV-MCK-3990 (Fig. 2a & 2b and Table 1). As expected, the equivalent muscle from the age-matched healthy C57/B10 mice showed indistinguishable dystrophin staining pattern, when stained with an antibody (Dys2) that recognizes both mouse and human dystrophin C-terminal region. However, this antibody (Dys2) failed to stain the AAV treated *mdx* muscle (data not shown) due to deletion of this region in our mini-dystrophin genes, confirming the identity of the vector-derived transgene products. Consistently, the untreated *mdx* control muscle showed no

dystrophin staining (Fig. 2) except the very few somatic revertant myofibers recognized by Dys2 antibody. We next examined whether the mini-dystrophins were functional in restoring the missing DAP complexes, including the sarcoglycan complex on the myofiber membrane which is not found in untreated dystrophic muscle. Immunofluorescent staining using three antibodies against α, β, and γ sarcoglycans respectively, showed positive results in all of the consecutive thin sections adjacent to those stained with dystrophin antibodies (Fig. 2). These results provided evidence of biochemical functionality of the mini-dystrophins (lacking the entire distal C-terminal region) in interactions with the DAP complexes.

Histological examination of the AAV mini-dystrophin treated muscles showed nearly exclusive (~99%) peripheral nucleation in the mini-dystrophin positive myofibers, as revealed by dystrophin immunostaining and myonuclei counterstaining with DAPI (Fig. 2 and Table 1). The mutual exclusivity between mini-dystrophin expression and central nucleation in the vector treated *mdx* muscle precisely mirrors that of the normal muscle (Table 1). In addition, the myofibers positive for mini-dystrophin expression also exhibited consistent myofiber sizes and polygonal shapes indistinguishable from those of the normal muscles (Fig. 2). By contrast, the untreated *mdx* muscle showed extensive central nucleation (Table 1), with additional signs of dystrophic pathology including wide variation of myofiber sizes, round myofiber shapes, and fibrosis (Fig.2). Hence, vector treatment eliminates dystrophic pathology and led to normal histology in terms of peripheral nucleation, consistent myofibre size and lack of fibrosis in the mini-dystrophin positive areas. These results unequivocally demonstrated the absence of muscle degeneration due to the therapeutic effects of the novel mini-dystrophins.

Plasma membrane damage and leakage in dystrophic muscle is a major physiological defect. In order to determine weather a functional mini-dystrophin would be effective in improving plasma membrane integrity of the dystrophic myofibers following AAV vector mediated gene therapy, muscle cell membrane integrity was examined using Evans Blue dye, a vital red-fluorescent dye that is excluded by the normal myofibers, but is taken up by the dystrophic myofibers with leaky cell membrane due to contractile damages. A previous study of *mdx* mice revealed that the apoptotic myonuclei were exclusively found in Evans Blue dye positive myofibers, correlating membrane leakage and muscle cell apoptosis. [Matsuda, R. et al.,

J Biochem (Tokyo) 118, 959-964 (1995)]. In our experiments, the dye was administered into the tail vein of 15-week old vector-treated and untreated *mdx* mice as well as age-matched healthy mice. To induce mechanical stress to the muscle, the mice were allowed to exercise by continuous swimming for 20 minutes. Fifteen hours after dye administration, muscle samples were collected and examined for dystrophin expression along with Evans Blue dye uptake. Muscle from healthy mice revealed no uptake of the dye by the myofibers and uniform dystrophin staining across the muscle sections (Fig. 3, first row). The AAV vector treated *mdx* muscle showed results consistent with the normal muscle, thus demonstrating the mini-dystrophins were sufficient to prevent plasma membrane damage (Fig. 3, second to fourth rows). Dye uptake (red fluorescence) was found only in those myofibers which stained negative for mini-dystrophin and confined in the area not transduced by the AAV vectors (Fig. 3, second to fourth rows). By contrast, the untreated *mdx* muscle revealed no presence of dystrophin but extensive dye uptake (Fig. 3, last row). These results proved the physiological functionality of the novel mini-dystrophins in maintaining membrane integrity and protecting myofibers from mechanical damages by muscle contraction.

In summary, the present invention provides that the dystrophin gene can be successfully reduced to one third (1/3) of its 11 kb full-length coding sequence, without compromising essential functions in protecting muscles from dystrophic phenotypes. Moreover, the present invention provides AAV vectors carrying the mini-genes are capable of mediating efficient and stable correction of both biochemical and physiological defects in a major muscle group of a DMD animal model. Previous attempts to generate mini-genes that were shorter than 1/2 of the full length dystrophin failed to preserve the essential protective functions. [Yuasa, K. *et al.*, *FEBS Lett* 425, 329-336 (1998)]. Although the mini-genes contained both intact N- and C-terminal domains and 1 to 3 central rod repeats, they were functionally similar to a C-terminal dystrophin construct (Dp71), [Cox, G. A. et al., *Nat Genet* 8, 333-339 (1994); Greenberg, D. S. et al., *Nat Genet* 8, 340-344 (1994)], and thus sufficient to restore DAP complexes but insufficient to protect muscle from dystrophic pathology. However, the mini-dystrophin genes reported here accommodated at least 5 rod repeats (R1, R2, R22, R23 and R24) and two hinges (H1 and H4). Therefore we hypothesized that the length of the central rod domain is the most

critical factor, based on the fact that a major role of dystrophin is to crosslink the myofiber cytoskelton and plasma membrane and stabilize the structure during muscle contraction. If the dystrophin is too short to span the sliding distance between the cytoskeleton and plasma membrane during muscle contraction, the crosslink will be disrupted and the muscle membrane will become unstable and prone to mechanical damages. To accommodate as many rod units in the central domain without exceeding the AAV vector packaging limit, we have for the first time deleted the entire C-terminus (819 bp) without sacrificing the primary functions of dystrophin. Our results indicate that 5 rods and 2 hinges seemed to provide sufficient length and flexibility for the central domain. This conclusion is supported by the observation that mini-genes Δ3849 and Δ3990 were equally functional in preventing the dystrophic phenotypes, although Δ3990 has an extra hinge (H3). Similarly, mini-gene Δ4173 has an extra rod (R3) but did not function better than mini-genes Δ3849 or Δ3990 (Table 1). In fact, because the entire AAV-MCK-4173 vector cassette is nearly 5.2 kb in length, larger than the 5 kb packaging limit, the infectivity of its viral particles was impaired leading to lower gene transfer efficiency (Fig. 3 and Table 1).

EXAMPLES

Example 1: Construction of mini-dystrophin genes and AAV vector production

Mini-dystrophin constructs were created mainly by PCR cloning method with Pfu polymerase (Stratagene, CA) from human dystrophin cDNA (GenBank # NM 004006). For consistency, the numbering of the nucleotide only includes the 11,058 bp dystrophin coding sequence. Mini-gene Δ 3849 contains nucleotides 1 to 1668, 8059 to 10227 and 11047 to 11058. Mini-gene Δ3990 contains nucleotides 1-1668, 7270-7410, 8059-10227 and 11047-11058. Mini-gene Δ4173 contains nucleotides 1-1992 and 8059-10227 and 11047-11058. The above constructs were subcloned into an AAV vector plasmid containing an MCK promoter[23] and a small polyA signal sequence (a gift from T. R. Flotte, University of Florida) to generate vector constructs AAV-MCK-3849, AAV-MCK-3990 and AAV-MCK-4173. AAV viral particles were produced

according to previously published methods. [Xiao, X. et al., *Journal of Virology* **72**, 2224-2232 (1998)].

Example 2. Animal and vector administration

All experiments involving animals were approved by the University of Pittsburgh Animal Care & Use Committee. The healthy mice *C57/B10* and dystrophic mice *mdx* were purchased from The Jackson Laboratory (Bar Harbor, Maine). The 10-day old *mdx* pups were injected into the hindleg gastrocnemius muscle with 50 µl (5 x 10^{10} viral particles) of AAV-MCK-3849, AAV-MCK-3990 or AAV-MCK-4173. Three months after vector injection, muscle samples were collected for examination.

Example 3. Immunofluorescent staining

Muscle cryosections of 5 µm thickness were immunofluorescently stained with the Mouse-on-Mouse Kit from the Vector Laboritories (Burlingame, CA) according to the manufacturer's protocol, except that the cryosections were immediately treated with the blocking buffer without the fixation step. Monoclonal antibodies against dystrophin (NCL-Dys3, N-terminal-specific and human-specific; NCL-Dys2, C-terminal-specific), and antibodies against α-, β-, and γ-sarcoglycans (NCL-a-SARC, NCL-b-SARC and NCL-g-SARC) were purchased from Novocastra Laboratories Ltd (Burlingame, CA). Cell nuclei were counter-stained with 0.01% DAPI (Sigma, St. Louis, MO) for 10 minutes. Photographs were taken with a Nikon TE-300 fluorescent microscope.

***In vivo* myofiber plasma membrane integrity test:** [Matsuda, R. et al., *J Biochem (Tokyo)* **118**, 959-964 (1995)] Evans Blue dye (10 mg/ml PBS) was injected into the tail vein of 15-week old C57/B10 mice, *mdx* mice, and AAV vector-treated *mdx* mice at the dose of 0.1 mg/gram of body weight. Following dye injection, mice were allowed continuous swimming for 20 minutes. At 15 hours after Evans Blue injection, the mice were sacrificed and muscle were cryosectioned. Evans

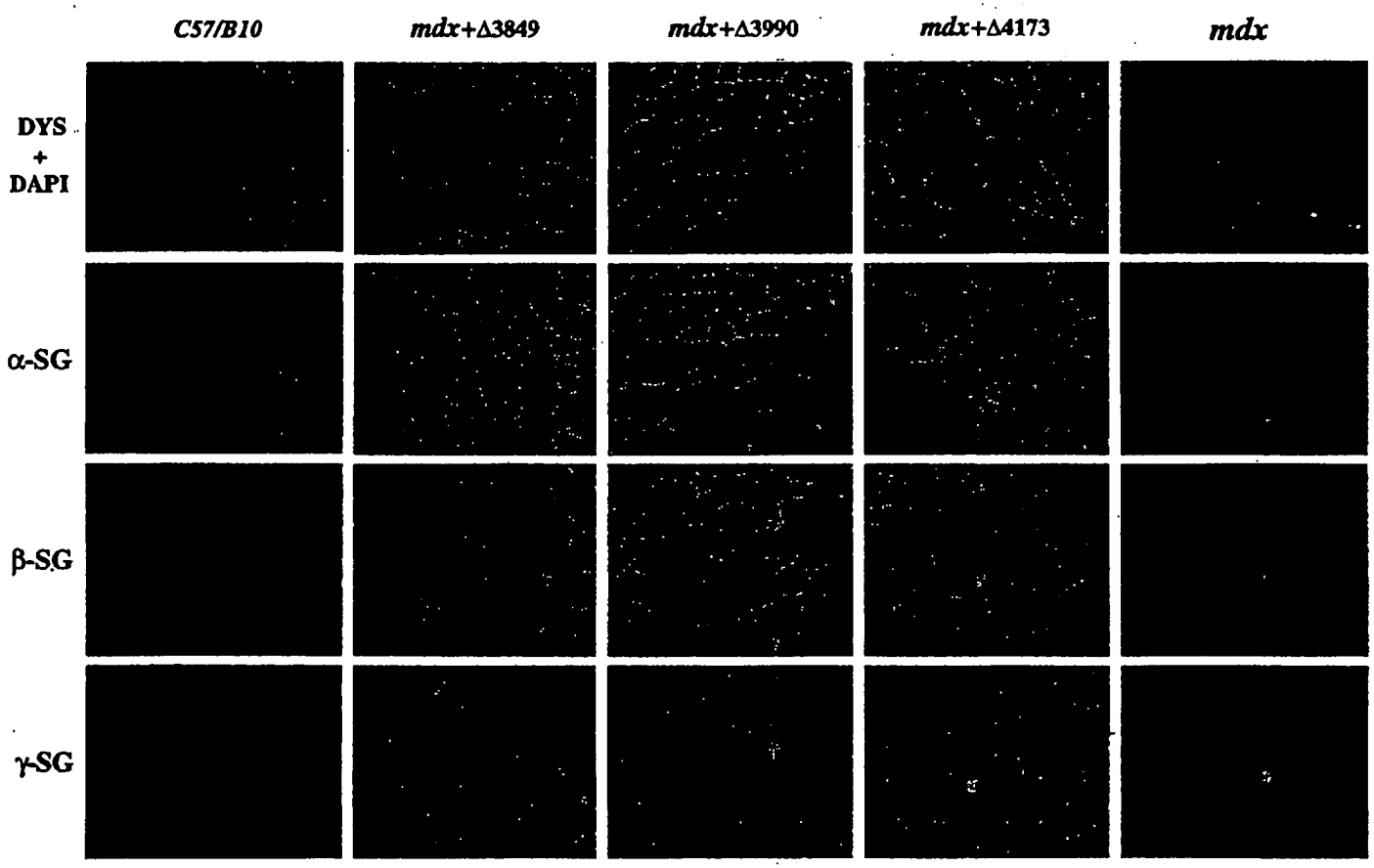
Table 1. In vivo AAV gene transfer efficiency and prevention of central nucleation

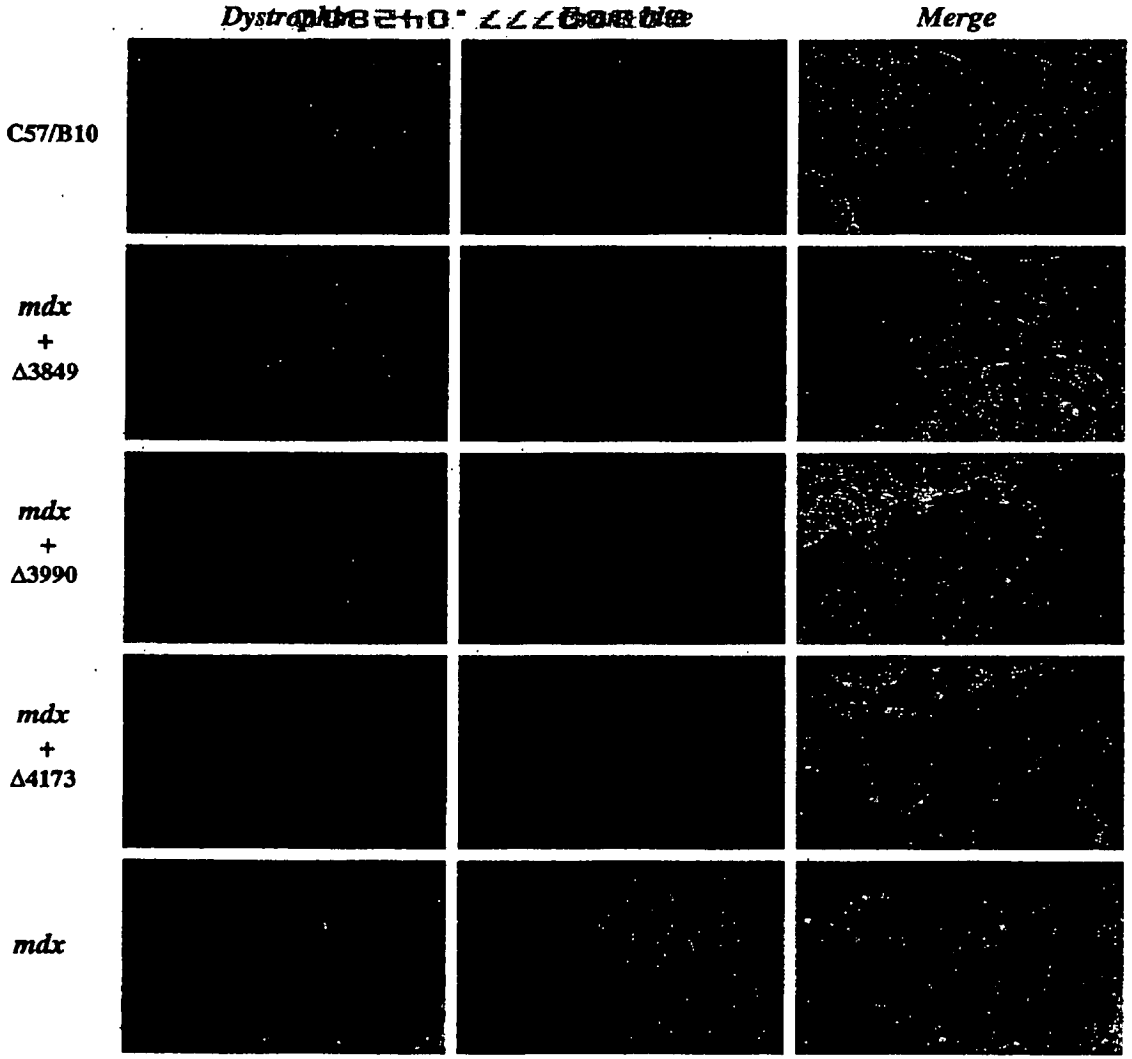
Animal* groups	n	Age at Vector injection	months post injection	% Dystrophin positive fibers	% Central nuclei **
C57/B10	4	No injection	N/A	100	1.45 (56/3860)
mdx+Δ3849	4	10 days	3	56%~88%	1.02 (72/7098)
mdx+Δ3990	4	10 days	3	50%~80%	0.99 (56/5652)
mdx+Δ4173	4	10 days	3	15%~25%	0.88 (41/4667)
mdx	4	No injection	N/A	< 1%	75.4 (2382/3160)

Note: * All animals were age-matched. N/A stands for not appliable.
** All numbers were collected from dystrophin positive myofibers, except that in *mdx* mice which were not injected with AAV vectors.

a

b




Dystrophin
Merge
C57/B10
mdx + Δ3849
mdx + Δ3990
mdx + Δ4173
mdx

VERIFIED STATEMENT (DECLARATION) CLAIMING SMALL ENTITY STATUS (37 CFR 1.9(f) AND 1.27 (b)) - INDEPENDENT INVENTOR	Docket No. 4268

Serial No.	Filing Date	Patent No.	Issue Date

Applicant/ Patentee: Xiao XIAO

Invention: ADENO-ASSOCIATED VIRAL VECTORS CARRYING NOVEL HUMAN MINI-DYSTROPHIN GENES

As a below named inventor, I hereby declare that I qualify as an independent inventor as defined in 37 CFR 1.9(c) for purposes of paying reduced fees under section 41(a) and (b) of Title 35, United States Code, to the Patent and Trademark Office with regard to the invention entitled above and described in:

- ☒ the specification to be filed herewith.
- ☐ the application identified above.
- ☐ the patent identified above.

I have not assigned, granted, conveyed or licensed and am under no obligation under contract or law to assign, grant, convey or license, any rights in the invention to any person who could not be classified as an independent inventor under 37 CFR 1.9(c) if that person had made the invention, or to any concern which would not qualify as a small business concern under 37 CFR 1.9(d) or a nonprofit organization under 37 CFR 1.9(e).

Each person, concern or organization to which I have assigned, granted, conveyed, or licensed or am under an obligation under contract or law to assign, grant, convey, or license any rights in the invention is listed below:

- ☒ No such person, concern or organization exists.
- ☐ Each such person, concern or organization is listed below.

*NOTE: Separate verified statements are required from each named person, concern or organization having rights to the invention averring to their status as small entities (37 CFR 1.27)

FULL NAME _____

ADDRESS _____

☐ Individual ☐ Small Business Concern ☐ Nonprofit Organization

FULL NAME _____

ADDRESS _____

☐ Individual ☐ Small Business Concern ☐ Nonprofit Organization

FULL NAME _____

ADDRESS _____

☐ Individual ☐ Small Business Concern ☐ Nonprofit Organization

FULL NAME _____

ADDRESS _____

☐ Individual ☐ Small Business Concern ☐ Nonprofit Organization

I acknowledge the duty to file, in this application or patent, notification of any change in status resulting in loss of entitlement to small entity status prior to paying, or at the time of paying, the earliest of the issue fee or any maintenance fee due after the date on which status as a small entity is no longer appropriate. (37 CFR 1.28(b))

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application, any patent issuing thereon, or any patent to which this verified statement is directed.

NAME OF INVENTOR Xiao XIAO
SIGNATURE OF INVENTOR _____ DATE: 04-27-00

NAME OF INVENTOR _____
SIGNATURE OF INVENTOR _____ DATE: _____

NAME OF INVENTOR _____
SIGNATURE OF INVENTOR _____ DATE: _____

NAME OF INVENTOR _____
SIGNATURE OF INVENTOR _____ DATE: _____

NAME OF INVENTOR _____
SIGNATURE OF INVENTOR _____ DATE: _____

NAME OF INVENTOR _____
SIGNATURE OF INVENTOR _____ DATE: _____

NAME OF INVENTOR _____
SIGNATURE OF INVENTOR _____ DATE: _____

NAME OF INVENTOR _____
SIGNATURE OF INVENTOR _____ DATE: _____

Please type a plus sign (+) inside this box [+]

05.07.00

A/PROV

Docket Number: 4268

PROVISIONAL APPLICATION FOR PATENT COVER SHEET (Small Entity)

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53 (c).

INVENTOR(S)/APPLICANT(S)		
Given Name (first and middle [if any])	Family Name or Surname	Residence (City and either State or Foreign Country)
Xiao	XIAO	2433 Dogwood Lane Wexford, PA 15090

[ ] Additional Inventors are being named on page 2 attached hereto

TITLE OF THE INVENTION (280 characters max)

ADENO-ASSOCIATED VIRAL VECTORS CARRYING NOVEL HUMAN MINI-DYSTROPHIN GENES

CORRESPONDENCE ADDRESS

Direct all correspondence to:

[ ] Customer Number [ ] → *Place Customer Number Bar Code Label here*

OR

[X] Firm or Individual Name	Anderson, Kill & Olick, P.C.				
Address	1251 Avenue of the Americas				
Address					
City	New York	State	New York	ZIP	10020-1182
Country	U.S.A.	Telephone	212-278-1000	Fax	212-278-1733

ENCLOSED APPLICATION PARTS (check all that apply)

[X]	Specification	*Number of Pages* 63	[X]	Small Entity Statement	
[X]	Drawing(s)	*Number of Sheets* 4	[ ]	Other (specify)	

METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT (check one)

		FILING FEE AMOUNT
[ ]	A check or money order is enclosed to cover the filing fees	
[X]	The Commissioner is hereby authorized to charge filing fees or credit any overpayment to Deposit Account Number: 01-1944	$75.00

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

[X] No.

[ ] Yes, the name of the U.S. Government agency and the Government contract number are: _____

Respectfully submitted,

SIGNATURE _____ Date April 28, 2000

TYPED or PRINTED NAME Richard B. Klar REGISTRATION NO. *(if appropriate)* 31,385

TELEPHONE _____

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

SEND TO: Box Provisional Application, Assistant Commissioner for Patents, Washington, DC 20231

P16SMALL/REV04

PROVISIONAL APPLICATION FOR PATENT COVER SHEET (Small Entity)

INVENTOR(S)/APPLICANT(S)		
Given Name (first and middle [if any])	Family Name or Surname	Residence (city and either State or Foreign Country)

Certificate of Mailing by Express Mail

I certify that this application and enclosed fee is being deposited on April 28, 2000 with the U.S. Postal Service "Express Mail Post Office to Addressee" service under 37 C.F.R. 1.10 and is addressed to the Assistant Commissioner for Patents, Washington, D.C. 20231.

Anne R. Jacoby

Signature of Person Mailing Correspondence

Anne R. Jacoby

Typed or Printed Name of Person Mailing Correspondence

EL179769475US

"Express Mail" Mailing Label Number

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

SEND TO: Box Provisional Application, Assistant Commissioner for Patents, Washington, DC 20231

EXPRESS MAIL CERTIFICATE OF MAILING - SEPARATE PAPER

IN THE MATTER OF: Xiao XIAO

ATTORNEY'S DOCKET NO.: 4268

FOR: ADENO-ASSOCIATED VIRAL VECTORS CARRYING NOVEL HUMAN MINI-DYSTROPHIN GENES

I hereby certify that the new provisional application with:

___X___ transmittal letter

___X___ specification and claims

_____ declaration

_____ assignment with Recordation Form Cover Sheet

___4___ sheets drawings

_____ priority documents

_____ preliminary amendment

___X___ Verified Statement claiming Small Entity status

_____

is being deposited with the United States Postal Services "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated below and is addressed to BOX PROVISIONAL APPLICATION Assistant Commissioner of Patents and Trademarks, Washington, DC 20231 on April 28, 2000

EXPRESS MAIL LABEL NO.

EL179769475US

Anne R. Jacoby
(Name of Person making deposit)

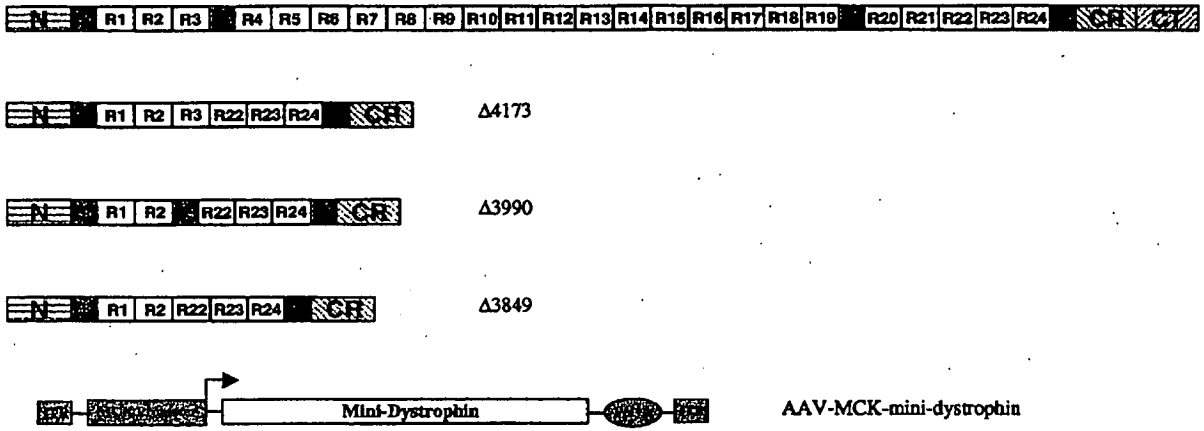
Anne R. Jacoby
(Signature)

April 28, 2000
(Date)

Anderson, Kill & Olick, P.C.
1251 Avenue of the Americas
New York, NY 10020-1182
(212)278-1000

60280777 042800

Human dystrophin coding sequence 11058bp



PTO/SB/31 (07-03)
Approved for use through 7/31/2003. OMB 0651-0031
U.S. Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

REQUEST FOR ACCESS TO AN ABANDONED APPLICATION UNDER 37 CFR 1.14

In re Application of

Application Number:	Filed:
60/200777	4-28-00

Paper No. #2

Bring completed form to:
File Information Unit
Crystal Plaza Three, Room 1D01
2021 South Clark Place
Arlington, VA
Telephone: (703) 308-2733

I hereby request access under 37 CFR 1.14(e)(1)(iv) to the application file record of the above-identified ABANDONED application, which is identified in, or to which a benefit is claimed, in the following document (as shown in the attachment):

United States Patent Application Publication No. 2003/0171312, page _____, line _____

United States Patent Number _____, column _____, line _____, or

WIPO Pub. No. _____; page _____, line _____

Related Information about Access to Pending Applications (37 CFR 1.14):

Direct access to pending applications is not available to the public but copies may be available and may be purchased from the Office of Public Records upon payment of the appropriate fee (37 CFR 1.19(b)), as follows:

For published applications that are still pending, a member of the public may obtain a copy of:
the file contents;
the pending application as originally filed; or
any document in the file of the pending application.

For unpublished applications that are still pending:

(1) If the benefit of the pending application is claimed under 35 U.S.C. 119(e), 120, 121, or 365 in another application that has: (a) issued as a U.S. patent, or (b) published as a statutory invention registration, a U.S. patent application publication, or an international patent application publication in accordance with PCT Article 21(2), a member of the public may obtain a copy of:
the file contents;
the pending application as originally filed; or
any document in the file of the pending application.

(2) If the application is incorporated by reference or otherwise identified in a U.S. patent, a statutory invention registration, a U.S. patent application publication, or an international patent application publication in accordance with PCT Article 21(2), a member of the public may obtain a copy of:
the pending application as originally filed.

Signature

5-5-04
Date

Typed or printed name

Registration Number, if applicable

(202) 415-1571
Telephone Number

FOR PTO USE ONLY

Approved by: _____ (Initials)

Unit: _____

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

10/06/00

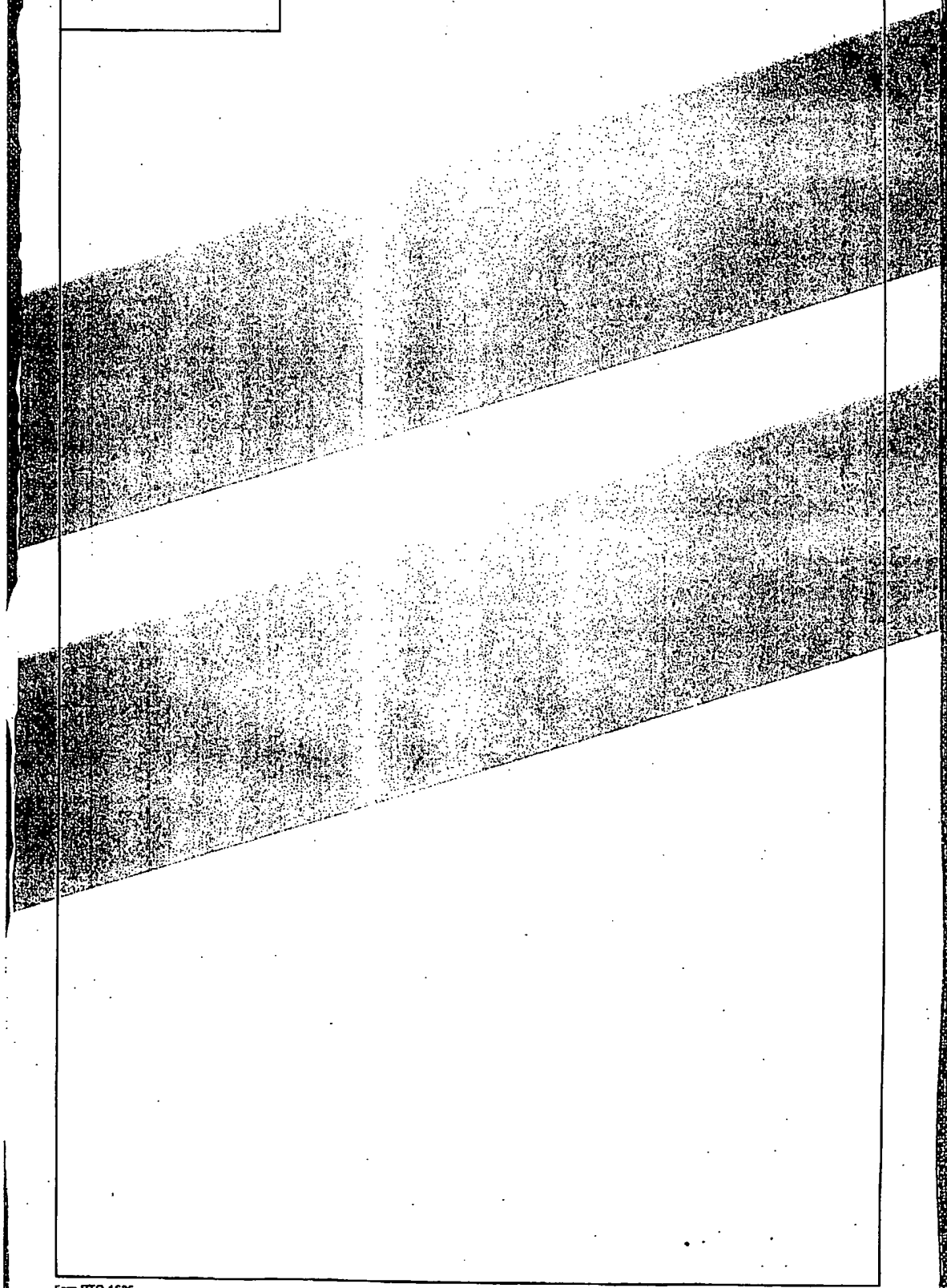
Class	Subclass

ISSUE CLASSIFICATION

[redacted]

art 49
QA CS4
VR

PROVISIONAL APPLICATION NUMBER



Form PTO-1625
(Rev. 6/99)

(FACE)

PATENT APPLICATION

60238848

APPROVED FOR LICENSE

INITIALS

CONTENTS

Date Entered or Counted		Date Received or Mailed
	1. Application ____ papers.	
	2. Request for Access	5/16/04
	3.	
	4.	
	5.	
	6.	
	7.	
	8.	
	9.	
	10.	
	11.	
	12.	
	13.	
	14.	
	15.	
	16.	
	17.	
	18.	
	19.	
	20.	
	21.	
	22.	
	23.	
	24.	
	25.	
	26.	
	27.	
	28.	
	29.	
	30.	
	31.	
	32.	

(LEFT OUTSIDE)

File History Report

While copying your file we noticed that there are one or more 'Other Publications' missing from the contents of the file wrapper. At your request, we will attempt to obtain the missing publication(s) from alternative resources. Please note that additional charges will apply to this service.

5350 Shawnee Road • Suite 110 • Alexandria, VA 22312
Phone: 800.445.9760 or 703.916.1500 • fax: 800.445.9761 or 703.916.1727
www.faxpat.com www.optipat.com

PATENT APPLICATION SERIAL NO. _____

U.S. DEPARTMENT OF COMMERCE
PATENT AND TRADEMARK OFFICE
FEE RECORD SHEET

10/17/2000 MYUSUF1 00000007 60238848
01 FC:114 150.00 OP

PTO-1556
(5/87)

*U.S. GPO: 1998-456-462/19144

UNITED STATES PATENT AND TRADEMARK OFFICE

COMMISSIONER FOR PATENTS
UNITED STATES PATENT AND TRADEMARK OFFICE
WASHINGTON, D.C. 20231
www.uspto.gov

Bib Data Sheet

SERIAL NUMBER	FILING DATE	CLASS	GROUP ART UNIT	ATTORNEY DOCKET NO.
60/238,848	10/06/2000 RULE -	-	-	UM-04723

APPLICANTS

Jeffrey S. Chamberlain, Ann Arbor, MI ;
Scott Q. Harper, Ann Arbor, MI ;

** CONTINUING DATA **********

** FOREIGN APPLICATIONS **********

IF REQUIRED, FOREIGN FILING LICENSE GRANTED ** 01/11/2001

Foreign Priority claimed	STATE OR COUNTRY	SHEETS DRAWING	TOTAL CLAIMS	INDEPENDENT CLAIMS
☐ yes ☐ no 35 USC 119 (a-d) conditions met ☐ yes ☐ no ☐ Met after Allowance Verified and Acknowledged Examiner's Signature Initials	MI	66	-	-

ADDRESS

Medlen & Carroll LLP
Suite 220
220 Montgomery Street
San Francisco ,CA 94104

TITLE

Truncated dystrophin genes

FILING FEE RECEIVED	FEES: Authority has been given in Paper	
150	No. _____ to charge/credit DEPOSIT ACCOUNT No. _____ for following:	☐ All Fees ☐ 1.16 Fees (Filing) ☐ 1.17 Fees (Processing Ext. of time) ☐ 1.18 Fees (Issue) ☐ Other _____ ☐ Credit

ABSTRACT

The present invention relates to compositions and methods for expressing mini-dystrophin peptides. In particular, the present invention provides compositions comprising nucleic acid sequences that are shorter than wild-type dystrophin cDNA and that express mini-dystrophin peptides that function in a similar manner as wild-type dystrophin proteins, and methods for expressing mini-dystrophin peptides in target cells. The present invention provides such shortened nucleic acid sequences in a variety of ways. For example, the present invention provides nucleic acid encoding only 4, 8, 12, 16, and 20 spectrin-like repeat encoding sequences (*i.e.* nucleic acid encoding an exact number of spectrin-like repeats that are multiples of 4). As wild-type dystrophin has 24 spectrin-like repeat encoding sequences, providing nucleic acid encoding fewer numbers of repeats reduces the size of the dystrophin gene (*e.g.* allowing the nucleic acid sequence to fit into vectors with limited cloning capacity). Another example of such shortened nucleic acid sequences are those that lack at least a portion of the carboxy-terminal domain of wild-type dystrophin nucleic acid. A further example of such shortened nucleic acid sequences are those that lack at least a portion of the 3' untranslated region, or 5' unstranslated region, or both.

TRUNCATED DYSTROPHIN GENES

This invention was made with Government support under contract NIH R01AR40864-10. The government has certain rights in this invention.

FIELD OF THE INVENTION

The present invention relates to compositions and methods for expressing mini-dystrophin peptides. In particular, the present invention provides compositions comprising nucleic acid sequences that are shorter than wild-type dystrophin cDNA and that express mini-dystrophin peptides that function in a similar manner as wild-type dystrophin proteins, and methods for expressing mini-dystrophin peptides in target cells.

BACKGROUND OF THE INVENTION

Muscular dystrophy is a group of inherited disorders characterized by progressive muscle weakness and loss of muscle tissue. Muscular dystrophies includes many inherited disorders, including Becker's muscular dystrophy and Duchenne's muscular dystrophy, which are both caused by mutations in the dystrophin gene. Both of the disorders have similar symptoms, although Becker's muscular dystrophy is a slower progressing form of the disease. Duchenne's muscular dystrophy is a rapidly progressive form of muscular dystrophy.

Both disorders are characterized by progressive muscle weakness of the legs and pelvis which is associated with a loss of muscle mass (wasting). Muscle weakness also occurs in the arms, neck, and other areas, but not as severely as in the lower half of the body. Calf muscles initially enlarge (an attempt by the body to compensate for loss of muscle strength), the enlarged muscle tissue is eventually replaced by fat and connective tissue (pseudohypertrophy). Muscle contractures occur in the legs and heels, causing inability to use the muscles because of shortening of muscle fibers and fibrosis of connective tissue. Bones develop abnormally, causing skeletal deformities of the chest and other areas. Cardiomyopathy occurs in almost all

cases. Mental retardation may accompany the disorder but it is not inevitable and does not worsen as the disorder progresses. The cause of this impairment is unknown.

Becker's muscular dystrophy occurs in approximately 3 out of 100,000 people. Symptoms usually appear in men between the ages of 7 and 26. Women rarely develop symptoms. There is no known cure for Becker's muscular dystrophy. Treatment is aimed at control of symptoms to maximize the quality of life. Activity is encouraged. Inactivity (such as bedrest) can worsen the muscle disease. Physical therapy may be helpful to maintain muscle strength. Orthopedic appliances such as braces and wheelchairs may improve mobility and self-care. Becker's muscular dystrophy results in slowly progressive disability. A normal life span is possible; however, death usually occurs after age 40.

Duchenne's muscular dystrophy occurs in approximately 2 out of 10,000 people. Symptoms usually appear in males 1 to 6 years old. Females are carriers of the gene for this disorder but rarely develop symptoms. There is no known cure for Duchenne's muscular dystrophy. Treatment is aimed at control of symptoms to maximize the quality of life. Activity is encouraged. Inactivity (such as bedrest) can worsen the muscle disease. Physical therapy may be helpful to maintain muscle strength and function. Orthopedic appliances such as braces and wheelchairs may improve mobility and the ability for self-care. Duchenne's muscular dystrophy results in rapidly progressive disability. By age 10, braces may be required for walking, and by age 12, most patients are confined to a wheelchair. Bones develop abnormally, causing skeletal deformities of the chest and other areas. Muscular weakness and skeletal deformities contribute to frequent breathing disorders. Cardiomyopathy occurs in almost all cases. Intellectual impairment is common but is not inevitable and does not worsen as the disorder progresses. Death usually occurs by age 15, typically from respiratory (lung) disorders.

Although there are no available treatments for muscular dystrophy, the usefulness of gene replacement as therapy for the disease has been established in transgenic mouse models. Unfortunately, progress toward therapy for human patients has been limited by lack of a suitable technique for delivery of such vectors to large

masses of muscle cells. What is needed in the art is a vector that can carry most of the dystrophin coding sequence, that can be cheaply produced in large quantities, that can be delivered to a large mass of muscle cells, and that provides stable expression of dystrophin after delivery.

SUMMARY OF THE INVENTION

The present invention relates to compositions and methods for expressing mini-dystrophin peptides. In particular, the present invention provides compositions comprising nucleic acid sequences that are shorter than full-length wild-type dystrophin cDNA and that express mini-dystrophin peptides that function in a similar manner as wild-type dystrophin proteins, and methods for expressing mini-dystrophin peptides in target cells. The present invention provides such shortened nucleic acid sequences in a variety of ways. For example, the present invention provides nucleic acids encoding only 4, 8, 12, 16, and 20 spectrin-like repeat encoding sequences (*i.e.* nucleic acids encoding an exact number of spectrin-like repeats that are multiples of 4). As wild-type dystrophin has 24 spectrin-like repeat encoding sequences, providing nucleic acids encoding fewer numbers of repeats reduces the size of the dystrophin gene (*e.g.* allowing the nucleic acid sequence to fit into vectors with limited cloning capacity). Another example of such shortened nucleic acid sequences are those that lack at least a portion of the carboxy-terminal domain of wild-type dystrophin nucleic acid. A further example of such shortened nucleic acid sequences are those that lack at least a portion of the 3' untranslated region, or 5' unstranslated region, or both.

In certain embodiments, the present invention provides compositions comprising nucleic acid encoding a mini-dystrophin peptide, wherein the mini-dystrophin peptide comprises a spectrin-like repeat domain, and wherein the spectrin-like repeat domain consists of n spectrin-like repeats, wherein n is an even number less than 24. In some embodiments, n is 20 or less. In other embodiments, n is 16 or less. In particular embodiments, n is 12 or less. In additional embodiments, n is 8 or less. In preferred embodiments, n is 4. In some embodiments, the present invention provides compositions comprising nucleic acid encoding a mini-dystrophin peptide,

wherein the mini-dystrophin peptide comprises a spectrin-like repeat domain, and wherein the spectrin-like repeat domain consists of n spectrin-like repeats, wherein n is 4, 8, 12, 16, or 20.

In certain embodiments, the mini-dystrophin peptide is capable of altering a measurable muscle value in a DMD animal model by at least approximately 10% of the wild type value. In other embodiments, the mini-dystrophin peptide is capable of altering a measurable muscle value in a DMD animal model by at least approximately 20% of the wild type value. In particular embodiments, the mini-dystrophin peptide is capable of altering a measurable muscle value in a DMD animal model by at least approximately 30% of the wild type value. In preferred embodiments, the mini-dystrophin peptide is capable of altering a measurable muscle value in a DMD animal model to a level similar to the wild-type value (*e.g.* ±4%). In certain embodiments, the nucleic acid comprises at least 2 spectrin-like repeat encoding sequences. In some embodiments, the spectrin-like repeat encoding sequences are precise spectrin-like repeat encoding sequences. In certain embodiments, the nucleic acid is less than 5 kilo-bases in length. In particular embodiments, the nucleic acid comprises viral DNA. In preferred embodiments, the viral DNA comprises adeno-associated viral DNA.

In certain embodiments, the present invention provides compositions comprising nucleic acid encoding a mini-dystrophin peptide, wherein the mini-dystrophin peptide comprises a spectrin-like repeat domain, and wherein the spectrin-like repeat domain consists of n spectrin-like repeats, wherein n is an even number less than 24; and wherein the nucleic acid comprises an actin-binding domain encoding sequence, a β-dystroglycan-binding domain encoding sequence, and at least 2 spectrin-like repeat encoding sequences. In some embodiments, the nucleic acid comprises at least 4 spectrin-like repeat encoding sequences.

In certain embodiments, the present invention provides compositions comprising nucleic acid, wherein the nucleic acid comprises at least 2 spectrin-like repeat encoding sequences, and wherein the nucleic acid encodes a mini-dystrophin peptide comprising a spectrin-like repeat domain, wherein the spectrin-like repeat

domain consists of n spectrin-like repeats, and wherein n is an even number less than 24. In some embodiments, the nucleic acid comprises at least 4 spectrin-like repeat encoding sequences.

In some embodiments, the nucleic acid comprises SEQ ID NO:39 (*i.e.* ΔR4-R23). In other embodiments, the nucleic acid comprises SEQ ID NO:40 (*i.e.* ΔR2-R21). In certain embodiments, the nucleic acid comprises SEQ ID NO:41 (*i.e.* ΔR2-R21+H3). In still other embodiments, the nucleic acid comprises SEQ ID NO:42 (*i.e.* ΔH2-R19).

In certain embodiments, the nucleic acid comprises an expression vector (*e.g.* plasmid, virus, etc). In some embodiments, the expression vector comprises viral DNA. In certain embodiments, the viral DNA comprises adeno-viral DNA. In some embodiments, the viral DNA comprises lentiviral DNA. In other embodiments, the viral DNA comprises helper-dependent adeno-viral DNA. In preferred embodiments, the viral DNA comprises adeno-associated viral DNA. In some embodiments, the nucleic acid is inserted in a virus (*e.g.* adeno-associated virus, adenovirus, helper-dependent adeno-associated virus, lentivirus).

In certain embodiments, the nucleic acid comprises an actin-binding domain encoding sequence. In particular embodiments, the actin binding domain comprises at least a portion of SEQ ID NO:6 (*e.g.* 5%, 10%, 20%, 40%, 50%, or 75% of SEQ ID NO:6). In other embodiments, the actin binding domain comprises at least a portion of a homolog or mutated version of SEQ ID NO:6 (*e.g.* 5%, 10%, 20%, 40%, 50%, or 75% of a SEQ ID NO:6 homolog or mutated version of SEQ ID NO:6). In certain embodiments, the nucleic acid comprises a β-dystroglycan binding domain. In certain embodiments, the β-dystroglycan binding domain comprises at least a portion of a dystrophin hinge 4 encoding sequence (*e.g.* the 3' 50% of SEQ ID NO:34), and at least a portion of dystrophin cysteine-rich domain encoding sequence (*e.g.* the 5' 75% of SEQ ID NO:35). In particular embodiments, at least a portion of hinge 4 is the WW domain (SEQ ID NO:45), or a homolog or mutation thereof.

In particular embodiments, the spectrin-like repeat encoding sequences are selected from the group consisting of SEQ ID NOS:8-10, 12-27, and 29-33. In some embodiments, the spectrin-like repeat encoding sequences are selected from the group consisting of SEQ ID NOS:8-10, 12-27, and 29-33, and homologs or mutations of SEQ ID NOS:8-10, 12-27, and 29-33. In preferred embodiments, the spectrin-like repeat encoding sequences are selected from the group consisting of SEQ ID NOS:8-10 and 29-33. In some embodiments, the spectrin-like repeat encoding sequences are identical (*e.g.* all the sequences are SEQ ID NO:8). In preferred embodiments, the spectrin-like repeat encoding sequences are all different (*e.g.* the nucleic acid sequence has only 4 spectrin-like repeat encoding sequences, and these 4 are: SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, and SEQ ID NO:33). In certain embodiments, nucleic acid sequence comprises at least one spectrin-like repeat encoding sequence selected from the group consisting of SEQ ID NOS:8-10, and at least one spectrin-like repeat encoding sequence selected from the group consisting of SEQ ID NOS:29-33.

In certain embodiments, the nucleic acid comprises at least one dystrophin hinge region. In some embodiments, the nucleic acid comprises at least one dystrophin hinge region selected from hinge region 1, hinge region 2, hinge region 3 and hinge region 4. In some embodiments, the nucleic acid comprises at least one dystrophin hinge region selected from hinge region 1, hinge region 2, and hinge region 3. In particular embodiments, dystrophin hinge region 1 is SEQ ID NO:7, or a homolog (*See, e.g.* Fig. 11), or a mutant version thereof. In particular embodiments, dystrophin hinge region 2 is SEQ ID NO:11, or a homolog (*See, e.g.* Fig. 11), or a mutant version thereof. In certain embodiments, dystrophin hinge region 3 is SEQ ID NO:28, or a homolog (*See, e.g.* Fig. 11), or a mutant version thereof. In other embodiments, dystrophin hinge region 4 is SEQ ID NO:34, or a homolog (*See, e.g.* Fig. 11), or a mutant version thereof.

In some embodiments, the nucleic acid comprises at least a portion of wild-type dystrophin C-terminal protein. In other embodiments, the nucleic acid comprises at least a portion of the 5' untranslated region. In particular embodiments, the nucleic

acid comprises at least a portion of the 3' untranslated region. In different embodiments, the nucleic acid sequence comprises regulatory sequences (*e.g.* MCK enhancer and promoter elements). In particular embodiments, the nucleic acid sequence is operably linked to regulatory sequences (*e.g.* MCK enhancer and promoter elements). In certain embodiments, the nucleic acid sequence comprises a mutant muscle-specific enhancer region.

In particular embodiments, the nucleic acid has less than 75% of a wild type dystrophin 5' untranslated region. In other embodiments, the nucleic acid has less than 50% or 20% or 1% (*e.g.* 0, 1, 2 nucleotides from a wild type dystrophin 5' untranslated region). In particularly preferred embodiments, the nucleic acid sequence does not contain any of the wild-type dystrophin 5' untranslated region. In certain embodiments, the nucleic acid has less than 75% of a wild type dystrophin 3' untranslated region. In other embodiments, the nucleic acid has less than 50%, preferably less than 40%, more preferably less than 35% of a wild type dystrophin 3' untranslated region. In certain embodiments, the nucleic acid does not contain a wild-type dystrophin 3' untranslated region (or, in some embodiments, any type of 3' untranslated region).

In particular embodiments, the mini-dystrophin peptide comprises a substantially deleted dystrophin C-terminal domain. In some embodiments, the mini-dystrophin peptide comprises less than 40% of wild type dystrophin C-terminal domain, preferably less than 30%, more preferably less than 20%, even more preferably less than 1%, and most preferably approximately 0% (*e.g.* 0, 1, 2, 3 or 4 amino acids from the wild type dystrophin C-terminal domain). In some embodiments, the nucleic acid sequence comprises at least one intron sequence.

In some embodiments, the present invention provides methods for expressing a mini-dystrophin peptide in a target cell, comprising; a) providing; i) a vector comprising nucleic acid encoding a mini-dystrophin peptide, wherein the mini-dystrophin peptide comprises a spectrin-like repeat domain, and wherein the spectrin-like repeat domain consists of n spectrin-like repeats, wherein n is an even number less than 24, and ii) a target cell, and b) contacting the vector with the target cell

under conditions such that the mini-dystrophin peptide is expressed in the target cells. In certain embodiments, the contacting comprises transfecting. In other embodiments, the target cell is a muscle cell. In particular embodiments, the target cell further comprises a subject (e.g. with Duchenne muscular dystrophy (DMD) or Becker muscular dystrophy (BMD)). In preferred embodiment, the mini-dystrophin peptide is expressed in the cells of a subject (*e.g.* such that symptoms of DMD or BMD are reduced or eliminated).

In particular embodiments, the present invention provides compositions comprising nucleic acid, wherein the nucleic acid encodes a mini-dystrophin peptide, and wherein the mini-dystrophin peptide comprises a substantially deleted dystrophin C-terminal domain. In certain embodiments, the substantially deleted dystrophin C-terminal domain is less than 40% of a wild type dystrophin C-terminal domain. In other embodiments, the substantially deleted dystrophin C-terminal domain is less than 30%, 20%, or 1% of a wild type dystrophin C-terminal domain. In preferred embodiments, the substantially deleted dystrophin C-terminal domain is approximately 0% of a wild type dystrophin C-terminal domain. In certain embodiments, the mini-dystrophin peptide does not contain any portion of the wild type dystrophin C-terminal domain (*i.e.* it is completely deleted).

In certain embodiments, the mini-dystrophin peptide is capable of altering a measurable muscle value in a DMD animal model by at least 10% of the wild type value. In other embodiments, the mini-dystrophin peptide is capable of altering a measurable muscle value in a DMD animal model by at least 20% of the wild type value. In particular embodiments, the mini-dystrophin-peptide is capable of altering a measurable muscle value in a DMD animal model by at least 30% of the wild type value. In preferred embodiments, the mini-dystrophin peptide is capable of altering a measurable muscle value in a DMD animal model to a level similar to the wild-type value (*e.g.* ±4%).

In certain embodiments, the nucleic acid comprises an expression vector (*e.g.* plasmid, virus, etc). In some embodiments, the expression vector comprises viral DNA. In certain embodiments, the viral DNA comprises adeno-viral DNA. In some

embodiments, the viral DNA comprises leniviral DNA. In other embodiments, the viral DNA comprises helper-dependent adeno-viral DNA. In preferred embodiments, the viral DNA comprises adeno-associated viral DNA. In some embodiments, the nucleic acid is inserted in a virus (*e.g.* adeno-associated virus, adenovirus, helper-dependent adeno-associated virus, lentivirus).

In certain embodiments, the nucleic acid comprises an actin-binding domain encoding sequence. In particular embodiments, the actin binding domain comprises at least a portion of SEQ ID NO:6 (*e.g.* 5%, 10%, 20%, 40%, 50%, or 75% of SEQ ID NO:6). In other embodiments, the actin binding domain comprises at least a portion of a homolog or mutated version of SEQ ID NO:6 (*e.g.* 5%, 10%, 20%, 40%, 50%, or 75% of a SEQ ID NO:6 homolog or mutated version of SEQ ID NO:6). In certain embodiments, the nucleic acid comprises a β-dystroglycan binding domain. In certain embodiments, the β-dystroglycan binding domain comprises at least a portion of a dystrophin hinge 4 encoding sequence (*e.g.* the 3' 50% of SEQ ID NO:34), and at least a portion of dystrophin cysteine-rich domain encoding sequence (e.g. the 5' 75% of SEQ ID NO:35). In particular embodiments, at least a portion of hinge 4 is the WW domain (SEQ ID NO:45), or a homolog or mutation thereof.

In certain embodiments, the nucleic acid comprises at least one dystrophin hinge region. In some embodiments, the nucleic acid comprises at least one dystrophin hinge region selected from hinge region 1, hinge region 2, hinge region 3 and hinge region 4. In some embodiments, the nucleic acid comprises at least one dystrophin hinge region selected from hinge region 1, hinge region 2, and hinge region 3. In particular embodiments, dystrophin hinge region 1 is SEQ ID NO:7, or a homolog (*See, e.g.* Fig. 11), or a mutant version thereof. In particular embodiments, dystrophin hinge region 2 is SEQ ID NO:11, or a homolog (*See, e.g.* Fig. 11), or a mutant version thereof. In certain embodiments, dystrophin hinge region 3 is SEQ ID NO:28, or a homolog (*See, e.g.* Fig. 11), or a mutant version thereof. In other embodiments, dystrophin hinge region 4 is SEQ ID NO:34, or a homolog (*See, e.g.* Fig. 11), or a mutant version thereof.

In other embodiments, the nucleic acid comprises at least a portion of the 5' untranslated region. In particular embodiments, the nucleic acid comprises at least a portion of the 3' untranslated region. In different embodiment, the nucleic acid sequence comprises regulatory sequences (*e.g.* MCK enhancer and promoter elements). In particular embodiments, the nucleic acid sequence is operably linked to regulatory sequences (*e.g.* MCK enhancer and promoter elements). In certain embodiments, the nucleic acid sequence comprises a mutant muscle-specific enhancer region.

In particular embodiments, the nucleic acid contains less that 75% of a wild type dystrophin 5' untranslated region. In other embodiments, the nucleic acid contains less than 50% or 20% or 1% (*e.g.* 0, 1, 2 nucleotides from a wild type dystrophin 5' untranslated region). In particularly preferred embodiments, the nucleic acid sequence does not contain any of the wild-type dystrophin 5' untranslated region. In certain embodiments, the nucleic acid has less than 75% of a wild type dystrophin 3' untranslated region. In other embodiments, the nucleic acid has less than 50%, preferably less than 40%, more preferably less than 35% of a wild type dystrophin 3' untranslated region. In certain embodiments, the nucleic acid does not contain a wild-type dystrophin 3' untranslated region (or, in some embodiments, any type of 3' untranslated region).

In some embodiments, the present invention provides methods for expressing a mini-dystrophin peptide in a target cell, comprising; a) providing; i) a vector comprising nucleic acid, wherein the nucleic acid encodes a mini-dystrophin peptide comprising a substantially deleted dystrophin C-terminal domain, and ii) a target cell, and b) contacting the vector with the target cell under conditions such that the mini-dystrophin peptide is expressed in the target cells. In certain embodiments, the contacting comprises transfecting. In other embodiments, the target cell is a muscle cell.

DESCRIPTION OF THE FIGURES

Figure 1 shows the nucleic acid sequence for wild-type human dystrophin cDNA.

Figure 2 shows the nucleic acid sequence for wild-type mouse dystrophin cDNA.

Figure 3 shows the nucleic acid sequence for wild-type human utrophin cDNA.

Figure 4 shows the nucleic acid sequence for wild-type mouse utrophin cDNA

Figure 5 shows various domains of the nucleic acid sequence for wild-type human dystrophin cDNA.

Figure 6 shows various domains of the nucleic acid sequence for wild-type human dystrophin cDNA.

Figure 7 shows various domains of the nucleic acid sequence for wild-type human dystrophin cDNA.

Figure 8 shows various domains of the nucleic acid sequence for wild-type human dystrophin cDNA.

Figure 9 shows various domains of the nucleic acid sequence for wild-type human dystrophin cDNA.

Figure 10 shows the 3' UTR domain nucleic acid sequence for wild-type human dystrophin cDNA.

Figure 11 shows a sequence alignment between wild-type human dystrophin cDNA and wild-type mouse dystrophin cDNA. The various domains in the human dystrophin sequence have spaces between them with the ends highlighted in bold. In this regard, homologous sequences for various domains in the mouse cDNA sequence are seen.

Figure 12 shows the nucleic acid sequence for ΔR4-R23, a nucleic acid sequence encoding a mini-dystrophin peptide.

Figure 13 shows the nucleic acid sequence for ΔR2-R21, a nucleic acid sequence encoding a mini-dystrophin peptide.

Figure 14 shows the nucleic acid sequence for ΔR2-R21+H3, a nucleic acid sequence encoding a mini-dystrophin peptide.

Figure 15 shows the nucleic acid sequence for ΔH2-R19, a nucleic acid sequence encoding a mini-dystrophin peptide.

Figure 16 shows the complete cDNA sequence for human skeletal muscle alpha actinin.

Figure 17 shows the nucleic acid sequence for ΔR9-R16, a nucleic acid sequence encoding a mini-dystrophin peptide.

Figure 18 shows the nucleic acid sequence for the WW domain.

Figure 19 shows various transgenic expression constructs tested in Example 1.

Figure 20 shows the contractile properties of EDL, soleus, and diaphragm muscles in wild-type, *mdx*, and dystrophin Δ71-78 mice.

Figure 21 show the nucleic acid sequence for pBSX.

Figure 22 shows a restriction map for pBSX.

Figure 23 shows the 'full-length' HDMD sequence.

Figure 24 shows the cloning procedure for ΔR4-R23.

Figure 25 shows the cloning procedure for ΔR2-R21+H3.

Figure 26 shows the cloning procedure for ΔR2-R21.

Figure 27 shows a schematic illustration of the domains encoded by the truncated and full-length dystrophins sequences tested in Example 5.

Figure 28 is a graph showing the percentage of myofibers in quadricep muscles of 3 month old mice that display centrally-located nuclei in the indicated strains of transgenic mice.

Figure 29 shows graphs depicting the force generating capacity in diaphragm (A) or EDL (B) muscles of the indicated strains of dystrophin transgenic *mdx* mice and control mice.

Figure 30 shows a graph depicting the force generating capacity in EDL (A) or diaphragm (B) muscles of the indicated strains of dystrophin transgenic *mdx* mice and control mice.

Figure 31 is a graph showing the percentage of force generating capacity lost after 1 or 2 lengthening contractions of the tibialis anterior muscle of the indicated strains of dystrophin transgenic *mdx* mice and control mice.

Figure 32 is a graph showing the total distance run on a treadmill by animals from the indicated strains of dystrophin transgenic *mdx* mice and control mice.

Figure 33 shows a graph depicting the total body mass (A) and mass of the tibialis anterior muscle (B) of the indicated strains of dystrophin transgenic *mdx* mice and control mice.

Figure 34 is a schematic illustration of the structure of a microdystrophin expression cassette inserted into an adeno-associated viral vector.

Figure 35 is a schematic illustration of the structure of plasmid pTZ19R (top) and the sequence of the multiple cloning site in the vector (bottom).

Figure 36 shows the nucleic acid sequence of various MCK enhancer regions (wild-type and mutant).

Figure 37 shows the nucleic acid sequence of various MCK promoter regions.

Figure 38 shows a comparison between domains in dystrophin and utrophin.

DEFINITIONS

To facilitate an understanding of the present invention, a number of terms and phrases are defined below:

As used herein, the term "measurable muscle values" refers to measurements of dystrophic symptoms (*e.g.* fibrosis, an increased proportion of centrally located nuclei, reduced force generation by skeletal muscle, etc.) in an animal. These measurements may be taken, for example, to determine the wild-type value (*i.e.* the value in a control animal), to determine the value in a DMD (Duchenne muscular dystrophy) animal model (*e.g.* in an *mdx* mouse model), and to determine the value in a DMD animal model expressing the mini-dystrophin peptides of the present invention. Various assays may be employed to determine measurable muscle values in an animal including, but not limited to, assays measuring fibrosis, phagocytic infiltration of muscle tissue, variation in myofiber size, an increased proportion of myofibers with centrally located nuclei, elevated serum levels of muscle pyruvate kinase, contractile

properties assays, DAP (dystrophin associated protein) assays, susceptibility to contraction induced injuries and measured force assays (*See* Examples 1 and 4).

As used herein, the term "mini-dystrophin peptide" refers to a peptide that is smaller in size than the full-length wild-type dystrophin peptide, and that is capable of altering (increasing or decreasing) a measurable muscle value in a DMD animal model by at least approximately 10% such that the value is closer to the wild-type value (*e.g.* a *mdx* mouse has a measurable muscle value that is 50% of the wild-type value, and this value is increased to at least 60% of the wild-type value; or a *mdx* mouse has a measurable muscle value that is 150% of the wild-type value, and this value is decreased to at most 140% of the wild-type value). In some embodiments, the mini-dystrophin-peptide is capable of altering a measurable muscle value in a DMD animal model by at least approximately 20% of the wild type value. In certain embodiments, the mini-dystrophin-peptide is capable of altering a measurable muscle value in a DMD animal model by at least approximately 30% of the wild type value. In preferred embodiments, the mini-dystrophin peptide is capable of altering a measurable muscle value in a DMD animal model to a level similar to the wild-type value (*e.g.* ±4%).

As used herein, the term "wild-type dystrophin cysteine-rich domain" refers to a peptide encoded by the nucleic acid sequences in SEQ ID NO:35 (*e.g.* in human), as well as wild type peptide homologs encoded by nucleic acid homologs of SEQ ID NO:35 (*See*, Fig. 11).

As used herein, the term "wild type dystrophin C-terminal domain" refers to a peptide encoded by the nucleic acid sequences in SEQ ID NO:36 (*e.g.* in human), as well as wild type peptide homologs encoded by nucleic acid homologs of SEQ ID NO:36 (*See*, Fig. 11).

As used herein, the term "mini-dystrophin peptide comprising a substantially deleted dystrophin C-terminal domain" refers to a mini-dystrophin peptide that has less than 45% of a wild type dystrophin C-terminal domain. In some embodiments, the mini-dystrophin peptide comprises less than 40% of wild type dystrophin C-terminal

domain, preferably less than 30%, more preferably less than 20%, even more preferably less than 1%, and most preferably approximately 0% (*e.g.* 0, 1, 2, 3 or 4 amino acids from the wild type dystrophin C-terminal domain). The construction of mini-dystrophin peptides with a substantially deleted dystrophin C-terminal domain may be accomplished, for example, by deleting all or a portion of SEQ ID NO:36 from human dystrophin SEQ ID NO:1 (*See, e.g.* Example 3C).

As used herein, the term "wild type dystrophin 5' untranslated region" refers to the nucleic acid sequence at the very 5' end of a wild type dystrophin nucleic acid sequence (*e.g.* SEQ ID NOS:1 and 2) that immediately precedes the amino acid coding regions. For example, for human dystrophin, SEQ ID NO:5 (the first 208 bases) is the 5' untranslated region (a homolog in mouse may be seen in Fig. 11).

As used herein, the term "wild type dystrophin 3' untranslated region" refers to the nucleic acid sequence at the very 3' end of a wild type dystrophin nucleic acid sequence (*e.g.* SEQ ID NOS:1 and 2) that immediately proceeds the amino acid coding regions. For example, for human dystrophin, SEQ ID NO:38 (the last 2690 bases of the human dystrophin gene) is the 3' untranslated region (a homolog in mouse may be seen in Fig. 11).

As used herein, the term "actin-binding domain encoding sequence" refers to the portion of a dystrophin nucleic sequence that encodes a peptide-domain capable of binding actin *in vitro* (*e.g.* SEQ ID NO:6), as well as homologs (*See*, Fig. 11), conservative mutations, and truncations of such sequences that encode peptide-domains that are capable of binding actin *in vivo*. Determining whether a particular nucleic acid sequence encodes a peptide-domain (*e.g.* homolog, mutation, or truncation of SEQ ID NO:6) that will bind actin *in vitro* may be performed, for example, by screening the ability of the peptide-domain to bind actin *in vitro* in a simple actin binding assay (*See*, Corrado *et al.*, FEBS Letters, 344:255-260 [1994], describing the expression of candidate dystrophin peptides as fusion proteins, absorbing F-actin on to microtiter plates, incubating the candidate peptides in the F-actin coated microtiter

plates, washing the plates, adding anti-fusion protein rabbit antibody, and adding an anti-rabbit antibody conjugated to a detectable marker).

As used herein, the term "β-dystroglycan-binding domain encoding sequence" refers to the portion of a dystrophin nucleic sequence that encodes a peptide-domain capable of binding β-dystroglycan *in vivo* (*e.g.* SEQ ID NOs:34 and 35), as well as homologs (*See*, Fig. 11), conservative mutations, and truncations of such sequences that encode peptide-domains that are capable of binding β-dystroglycan *in vivo*. In preferred embodiments, the β-dystroglycan-binding domain encoding sequence includes at least a portion of a hinge 4 encoding region (*e.g.* SEQ ID NO:45, the WW domain) and at least a portion of a wild-type dystrophin cysteine-rich domain (*e.g.* at least a portion of SEQ ID NO:35) (*See, e.g.* Jung *et al.*, *JBC*, 270 (45):27305 [1995]). Determining whether a particular nucleic acid sequence encodes a peptide-domain (*e.g.* homolog, mutation, or truncation) that will bind β-dystroglycan *in vivo* may be performed, for example, by first screening the ability of the peptide-domain to bind β-dystroglycan *in vitro* in a simple β-dystroglycan binding assay (*See*, Jung *et al.*, pg 27306 - constructing peptide-domain dystrophin-GST fusion peptides and radioactively labelled β-dystroglycan, immobilizing the fusion proteins on glutathione-agarose beads, incubating the beads with the radioactively labelled β-dystroglycan, pelleting the beads, washing the beads, and resolving the sample on an SDS-polyacrylamide gel, staining with Coomasie blue, exposing to film, and quantifying the amount of radioactivity present). Nucleic acid sequences found to express peptides capable of binding β-dystroglycan in such assays may then, for example, be tested *in vivo* by transfecting a cell line (*e.g.*, COS cells) with two expression vectors, one expressing the dystroglycan peptide and the other expressing the candidate peptide domain (as a fusion protein). After culturing the cells, the protein is then extracted and a co-immunoprecipitation is performed for one of the proteins, followed by a Western blot for the other.

As used herein, the term "spectrin-like repeats" refers to peptides composed of approximately 100 amino acids that are responsible for the rod-like shape of many

structural proteins including, but not limited to, dystrophin, utrophin, fodrin, alpha-actin, and spectrin, when the spectrin-like repeats are present in multiple copies (*e.g.* dystrophin-24, utrophin-22, alpha-actin-4, spectrin-16, ect). Spectrin-like repeats also refers to mutations of these natural peptides, such as conservative changes in amino acid sequence, as well as the addition or deletion of up to 5 amino acids to/from the end of a spectrin-like repeat. Spectrin-like repeats includes 'precise spectrin-like repeats' (see below). Examples of spectrin-like repeats include, but are not limited to, peptides encoded by nucleic acid sequences found in wild-type human dystrophin (*e.g.* SEQ ID NOS:8-10, 12-27, and 29-33).

As used herein, the term "spectrin-like repeat encoding sequences" refers to nucleic acid sequences encoding spectrin-like repeat peptides. This term includes natural and synthetic nucleic acid sequences encoding the spectrin-like repeats (*e.g.* both the naturally occurring and mutated spectrin-like repeat peptides). Examples of spectrin-like repeat encoding sequences include, but are not limited to, SEQ ID NOS:8-10, 12-27, and 29-33.

As used herein, the term "precise spectrin-like repeat encoding sequences" refers to nucleic acid sequences encoding spectrin-like repeat peptides with up to 1 additional amino acid added to, or deleted from, the spectrin-like repeat.

As used herein, the term "spectrin-like repeat domain" refers to the region in a mini-dystrophin peptide that contains the spectrin-like repeats of the mini-dystrophin peptide.

The term "gene" refers to a DNA sequence that comprises control and coding sequences necessary for the production of a polypeptide or precursor thereof. The polypeptide can be encoded by a full length coding sequence or by any portion of the coding sequence so long as the desired enzymatic activity is retained. The term "gene" encompasses both cDNA and genomic forms of a given gene.

The term "wild-type" refers to a gene, gene product, or other sequence that has the characteristics of that gene or gene product when isolated from a naturally occurring source. A wild-type gene is that which is most frequently observed in a population and is thus arbitrarily designated the "normal" or "wild-type" form of the

gene. In contrast, the term "modified" or "mutant" refers to a gene, gene product, or other sequence that displays modifications in sequence and or functional properties (*e.g.* altered characteristics) when compared to the wild-type gene or gene product. It is noted that naturally-occurring mutants can be isolated; these are identified by the fact that they have altered characteristics when compared to the wild-type gene or gene product.

The term "oligonucleotide" as used herein is defined as a molecule comprised of two or more deoxyribonucleotides or ribonucleotide, usually more than three (3), and typically more than ten (10) and up to one hundred (100) or more (although preferably between twenty and thirty). The exact size will depend on many factors, which in turn depends on the ultimate function or use of the oligonucleotide. The oligonucleotide may be generated in any manner, including chemical synthesis, DNA replication, reverse transcription, or a combination thereof.

As used herein, the term "regulatory sequence" refers to a genetic sequence or element that controls some aspect of the expression of nucleic acid sequences. For example, a promoter is a regulatory element that facilitates the initiation of transcription of an operably linked coding region. Other regulatory elements are enhancers, splicing signals, polyadenylation signals, termination signals, *etc.* Examples include, but are not limited to, the 5' UTR of the dystrophin gene (SEQ ID NO:5), MCK promotors and enhancers (both wild type and mutant, *See* U.S. provisional app. ser no. 60/218,436, hereby incorporated by reference).

Transcriptional control signals in eucaryotes comprise "promoter" and "enhancer" elements. Promoters and enhancers consist of short arrays of DNA sequences that interact specifically with cellular proteins involved in transcription. The present invention contemplates modified enhancer regions.

The term "recombinant DNA vector" as used herein refers to DNA sequences containing a desired coding sequence and appropriate DNA sequences necessary for the expression of the operably linked coding sequence in a particular host organism (*e.g.*, mammal). DNA sequences necessary for expression in procaryotes include a promoter, optionally an operator sequence, a ribosome binding site and possibly other

sequences. Eukaryotic cells are known to utilize promoters, polyadenlyation signals and enhancers.

The terms "in operable combination", "in operable order" and "operably linked" as used herein refer to the linkage of nucleic acid sequences in such a manner that a nucleic acid molecule capable of directing the transcription of a given gene and/or the synthesis of a desired protein molecule is produced. The term also refers to the linkage of amino acid sequences in such a manner so that a functional protein is produced.

"Hybridization" methods involve the annealing of a complementary sequence to the target nucleic acid (the sequence to be detected). The ability of two polymers of nucleic acid containing complementary sequences to find each other and anneal through base pairing interaction is a well-recognized phenomenon.

The "complement" of a nucleic acid sequence as used herein refers to an oligonucleotide which, when aligned with the nucleic acid sequence such that the 5' end of one sequence is paired with the 3' end of the other, is in "antiparallel association." Complementarity need not be perfect; stable duplexes may contain mismatched base pairs or unmatched bases. Those skilled in the art of nucleic acid technology can determine duplex stability empirically considering a number of variables including, for example, the length of the oligonucleotide, base composition and sequence of the oligonucleotide, ionic strength and incidence of mismatched base pairs.

The term "homology" refers to a degree of complementarity. There may be partial homology or complete homology (*i.e.*, identity). A partially complementary sequence is one that at least partially inhibits a completely complementary sequence from hybridizing to a target nucleic acid is referred to using the functional term "substantially homologous." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or Northern blot, solution hybridization and the like) under conditions of low stringency. A substantially homologous sequence or probe will compete for and inhibit the binding (*i.e.*, the hybridization) of a completely

homologous to a target under conditions of low stringency. This is not to say that conditions of low stringency are such that non-specific binding is permitted; low stringency conditions require that the binding of two sequences to one another be a specific (*i.e.*, selective) interaction. The absence of non-specific binding may be tested by the use of a second target that lacks even a partial degree of complementarity (*e.g.*, less than about 30% identity); in the absence of non-specific binding the probe will not hybridize to the second non-complementary target.

As used herein the term "stringency" is used in reference to the conditions of temperature, ionic strength, and the presence of other compounds such as organic solvents, under which nucleic acid hybridizations are conducted. Those skilled in the art will recognize that "stringency" conditions may be altered by varying the parameters just described either individually or in concert. With "high stringency" conditions, nucleic acid base pairing will occur only between nucleic acid fragments that have a high frequency of complementary base sequences (*e.g.*, hybridization under "high stringency" conditions may occur between homologs with about 85-100% identity, preferably about 70-100% identity). With medium stringency conditions, nucleic acid base pairing will occur between nucleic acids with an intermediate frequency of complementary base sequences (*e.g.*, hybridization under "medium stringency" conditions may occur between homologs with about 50-70% identity). Thus, conditions of "weak" or "low" stringency are often required with nucleic acids that are derived from organisms that are genetically diverse, as the frequency of complementary sequences is usually less.

Low stringency conditions when used in reference to nucleic acid hybridization comprise conditions equivalent to binding or hybridization at 42°C in a solution consisting of 5X SSPE (43.8 g/l NaCL, 6.9 g/l NaH_2PO_4-H_2O and 1.85 g/l EDTA, pH adjusted to 7.4 with NaOH), 0.1% SDA, 5X Denhardt's reagent [50X Denhardt's contains per 500 ml: 5 g Ficoll (Type 400, Pharmacia), 5 g BSA (Fraction V, Sigma)] and 100 µg/ml denatured salmon sperm DNA followed by washing in solution comprising 5X SSPE, 0.1% SDS at 42°C when a probe of about 500 nucleotides in length is employed.

High stringency condistions when used in reference to nucleic acid hybridization comprises conditions equivalent to binding or hybridizing at 42°C in a solution consisting of 5X SSPE (43.8 g/l NaCL, 6.9 g/l NaH_2PO_4-H_2O and 1.85 g/l EDTA, pH adjusted to 7.4 with NaOH), 0.5% SDS, 5X Denhardt's reagent and 100 ug/ml denatured salmon sperm DNA, followed by washing in a solution comprising 0.1X SSPE, 1.0% SDS at 42°C when a probe of about 500 nucleotides is employed.

The art knows well that numerous equivalent conditions may be employed to comprise low stringency conditions; factors such as the length and nature (DNA, RNA, base composition) of the probe and nature of the target (DNA, RNA, base composition, present in solution or immobilized, etc.) and the concentration of the salts and other components (*e.g.*, the presence or absence of formamide, dextran sulfate, polyethylene glycol) are considered and the hybridization solution may be varied to generate conditions of low stringency hybridization different from, but equivalent to, the above listed conditions. In addition, the art knows conditions that promote hybridization under conditions of high stringency (*e.g.*, increasing the temperature of the hybridization and/or wash steps, the use of formamide in the hybridization solution, etc.).

The term "transfection" as used herein refers to the introduction of foreign DNA into eukaryotic cells. Transfection may be accomplished by a variety of means known to the art including calcium phosphate-DNA co-precipitation, DEAE-dextran-mediated transfection, polybrene-mediated transfection, electroporation, microinjection, liposome fusion, lipofection, protoplast fusion, retroviral infection, and biolistics.

The term "stable transfection" or "stably transfected" refers to the introduction and integration of foreign DNA into the genome of the transfected cell. The term "stable transfectant" refers to a cell which has stably integrated foreign DNA into the genomic DNA.

As used herein, the terms "nucleic acid molecule encoding," "DNA sequence encoding," and "DNA encoding" refer to the order or sequence of deoxyribonucleotides along a strand of deoxyribonucleic acid. The order of these

deoxyribonucleotides determines the order of amino acids along the polypeptide (protein) chain. The DNA sequence thus codes for the amino acid sequence.

As used herein, the terms "muscle cell" refers to a cell derived from muscle tissue, including, but not limited to, cells derived from skeletal muscle, smooth muscle (*e.g.* from the digestive tract, urinary bladder, and blood vessels), and cardiac muscle. The term includes muscle cells *in vitro, ex vivo,* and *in vivo*. Thus, for example, an isolated cardiomyocyte would constitute a muscle cell, as would a cell as it exists in muscle tissue present in a subject *in vivo*. This term also encompasses both terminally differentiated and nondifferentiated muscle cells, such as myocytes, myotubes, myoblasts, cardiomyocytes, and cardiomyoblasts.

As used herein, the term "muscle-specific" in reference to an regulatory element (*e.g.* enhancer region, promoter region) means that the transcriptional activity driven by these regions is mostly in muscle cells or tissue (*e.g.* 20:1) compared to the activity conferred by the regulatory sequences in other tissues. An assay to determine the muscle-specificity of a regulatory region is provided in Example 5 below (measuring beta-galactoside in muscle cells and liver cells from a mouse transfected with an expression vector).

As used herein, the term "mutant muscle-specific enhancer region" refers to a wild-type muscle-specific enhancer region that has been modified (*e.g.* deletion, insertion, addition, substitution), and in particular, has been modified to contain an additional MCK-R control element (*See* U.S. Prov. App. Ser. No. 60/218,436, hereby incorporated by reference, and section IV below).

DESCRIPTION OF THE INVENTION

The present invention relates to compositions and methods for expressing mini-dystrophin peptides. In particular, the present invention provides compositions comprising nucleic acid sequences that are shorter than wild-type dystrophin cDNA and that express mini-dystrophin peptides that function in a similar manner as wild-type dystrophin proteins, and methods for expressing mini-dystrophin peptides in target

cells. The present invention provides such shortened nucleic acid sequences in a variety of ways. For example, the present invention provides nucleic acid encoding only 4, 8, 12, 16, and 20 spectrin-like repeat encoding sequences (*i.e.* nucleic acid encoding an exact number of spectrin-like repeats that are multiples of 4). As wild-type dystrophin has 24 spectrin-like repeat encoding sequences, providing nucleic acid encoding fewer numbers of repeats reduces the size of the dystrophin gene (*e.g.* allowing the nucleic acid sequence to fit into vectors with limited cloning capacity). Another example of such shortened nucleic acid sequences are those that lack at least a portion of the carboxy-terminal domain of wild-type dystrophin nucleic acid. A further example of such shortened nucleic acid sequences are those that lack at least a portion of the 3' untranslated region, or 5' unstranslated region, or both.

I. Dystrophin

A. Dystrophin Structure

In some embodiments, the present invention provides gene constructs comprising spectrin-like repeats from human dystrophin. Dystrophin is a 427 kDa cytoskeletal protein and is a member of the spectrin/αactinin superfamily (*See e.g.*, Blake *et al.*, Brain Pathology, 6:37 [1996]; Winder, J. Muscle Res. Cell. Motil., 18:617 [1997]; and Tinsley *et al.*, PNAS, 91:8307 [1994]). The N-terminus of dystrophin binds to actin, having a higher affinity for non-muscle actin than for sarcomeric actin. Dystrophin is involved in the submembraneous network of non-muscle actin underlying the plasma membrane. Dystrophin is associated with an oligomeric, membrane spanning complex of proteins and glycoproteins, the dystrophin-associated protein complex (DPC). The N-terminus of dystrophin has been shown *in vitro* to contain a functional actin-binding domain. The C-terminus of dystrophin binds to the cytoplasmic tail of β-dystroglycan, and in concert with actin, anchors dystrophin to the sarcolemma. Also bound to the C-terminus of dystrophin are the cytoplasmic members of the DPC. Dystrophin thereby provides a link between the actin-based

cytoskeleton of the muscle fiber and the extracellular matrix. It is this link that is disrupted in muscular dystrophy.

The central rod domain of dystrophin is composed of a series of 24 weakly repeating units of approximately 110 amino acids, similar to those found in spectrin (*i.e.*, spectrin-like repeats). This domain constitutes the majority of dystrophin and gives dystrophin a flexible rod-like structure. The rod-domain is interrupted by four hinge regions that are rich in proline. It is contemplated that the rod-domain provides a structural link between member of the DPC. Table 1 shows an overview of the structural and functional domains of human dystrophin.

Table 1 - Full Length Human Dystrophin cDNA

Nucleotides	**Feature**	**SEQ ID NO:**
1-208	5' untranslated region	SEQ ID NO:5
209-211	Start codon (ATG)	---
209-964	N terminus	SEQ ID NO:6
965-1219	Hinge 1	SEQ ID NO:7
1220-1546	Spectrin-like repeat No. 1	SEQ ID NO:8
1547-1879	Spectrin-like repeat No. 2	SEQ ID NO:9
1880-2212	Spectrin-like repeat No. 3	SEQ ID NO:10
2213-2359	Hinge 2	SEQ ID NO:11
2360-2692	Spectrin-like repeat No. 4	SEQ ID NO:12
2693-3019	Spectrin-like repeat No. 5	SEQ ID NO:13
3020-3346	Spectrin-like repeat No. 6	SEQ ID NO:14
3347-3673	Spectrin-like repeat No. 7	SEQ ID NO:15
3674-4000	Spectrin-like repeat No. 8	SEQ ID NO:16
4001-4312	Spectrin-like repeat No. 9	SEQ ID NO:17
4313-4588	Spectrin-like repeat No. 10	SEQ ID NO:18
4589-4915	Spectrin-like repeat No. 11	SEQ ID NO:19

4916-5239	Spectrin-like repeat No. 12	SEQ ID NO:20
5340-5551	Spectrin-like repeat No. 13	SEQ ID NO:21
5552-5833	Spectrin-like repeat No. 14	SEQ ID NO:22
5834-6127	Spectrin-like repeat No. 15	SEQ ID NO:23
6128-6187	20 amino acid insert (not hinge)	---
6188-6514	Spectrin-like repeat No. 16	SEQ ID NO:24
6515-6835	Spectrin-like repeat No. 17	SEQ ID NO:25
6836-7186	Spectrin-like repeat No. 18	SEQ ID NO:26
7187-7489	Spectrin-like repeat No. 19	SEQ ID NO:27
7490-7612	Hinge 3	SEQ ID NO:28
7613-7942	Spectrin-like repeat No. 20	SEQ ID NO:29
7943-8269	Spectrin-like repeat No. 21	SEQ ID NO:30
8270-8617	Spectrin-like repeat No. 22	SEQ ID NO:31
8618-9004	Spectrin-like repeat No. 23	SEQ ID NO:32
9005-9328	Spectrin-like repeat No. 24	SEQ ID NO:33
9329-9544	Hinge 4	SEQ ID NO:34
9545-10431	Start of C terminus	SEQ ID NO:35
10432-11254	Alternatively spliced exons 71-78	SEQ ID NO:36
11255-11266	End of Coding Region	SEQ ID NO:37
11267-13957	3' untranslated region	SEQ ID NO:38

* Domain structure based on Winder *et al., Febs Letters*, 369:27-33 (1995)

B. Spectrin-Like Repeats

Spectrin-like repeats are about 100 amino acids long and are found in a number of proteins, including the actin binding proteins spectrin, fodrin, α-actinin, and dystrophin, but their function remains unclear (Dhermy, 1991. Biol. Cell, 71:249-254). These domains may be involved in connecting functional domains and/or mediate protein-protein interactions. The many tandem, spectrin-like motifs that comprise most of the mass of the proteins in this superfamily are responsible for their similar flexible,

rod-like molecular shapes. Although these homologous motifs are frequently called repeats or repetitive segments, adjacent segments in each protein are only distantly related evolutionarily.

Spectrin is a cytoskeletal protein of red blood cells that is associated with the cytoplasmic side of the lipid bilayer (*See e.g.*, Speicher and Ursitti, Current Biology, 4:154 [1994]). Spectrin is a long-thin flexible rod-shaped protein that constitutes about 25% of the membrane-associated protein mass. Spectrin is composed of two large polypeptide chains, α-spectrin (~240 kDa) and β-spectrin (~220 kDa) and serves to cross-link short actin oligomers to form a dynamic two-dimensional submembrane latticework. Spectrin isoforms have been found in numerous cell types and have been implicated in a variety of functions.

The recent determination of the crystal structure of a single domain of spectrin provides insight into the structure function of an entire class of large actin cross-linking proteins (Yan et al., Science, 262:2027 [1993]). The domain is an example of a spectrin-like repeat. Early analysis of spectrin-like repeats by partial peptide sequence analysis demonstrated that most of the antiparallel spectrin heterodimer is made up of homologous 106 residue motifs. Subsequent sequence analyses of cDNAs confirmed that this small motif is the major building block for all spectrin isoforms, as well as for the related actinins and dystrophins (Matsudaira, Trends Biochem Sci, 16:87 [1991]).

Given their similar sequences, all spectrin motifs are expected to have related, but not identical, three-dimensional structures. The structure of a single Drosophila spectrin motif, 14, which has now been determined (Yan et al., Science, 262:2027 [1993]), should therefore provide insight into the overall conformation of spectrins in particular and, to a more limited extent, the other members of the spectrin superfamily. The structure shows that the spectrin motif forms a three-helix bundle, similar to the earliest conformational prediction based on the analysis of multiple homologous motifs (Speicher and Marchesi, Nature, 311:177 [1984]).

II. Variants and Homologs of Dystrophin

The present invention is not limited to the spectrin-like repeat encoding sequences SEQ ID NOS:8-10, 12-27, and 29-33, but specifically includes nucleic acid sequences capable of hybridizing to the spectrin-like repeat encoding sequences SEQ ID NOS::8-10, 12-27, and 29-33, (*e.g.* capable of hybridizing under high stringent conditions). Those skilled in the art know that different hybridization stringencies may be desirable. For example, whereas higher stringencies may be preferred to reduce or eliminate non-specific binding between the spectrin-like repeat encoding sequences SEQ ID NOS:8-10, 12-27, and 29-33, and other nucleic acid sequences, lower stringencies may be preferred to detect a larger number of nucleic acid sequences having different homologies to the nucleotide sequence of SEQ ID NOS:8-10, 12-27, and 29-33.

Accordingly, In some embodiments, the dystrophin spectrin-like repeats of the compositions of the present invention (*e.g.*, SEQ ID NOs:8-10, 12-27, and 29-33) are replaced with different spectrin-like repeats, including, but not limited to, variants, homologs, truncations, and additions of dystrophin spectrin-like repeats. Candidate spectrin-like repeats are screened for activity using any suitable assay, including, but not limited to, those described below and in illustrative Examples 1 and 5.

A. Homologs

1. Dystrophin From other Species

In some embodiments, the spectrin-like repeats of the gene constructs of the present invention are replaced with spectrin-like repeats of dystrophin from other species (*e.g.*, homologs of dystrophin), including, but not limited to, those described herein. Homologs of dystrophin have been identified in a variety of organisms, including mouse (Genbank accession number M68859); dog (Genbank accession number AF070485); and chicken (Genbank accession number X13369). The spectrin-like repeats of the mouse dystrophin gene were compared to the human gene (See Figure 11) and were shown to have significant homology. Similar comparisons can be

generated with homologs from other species, including but not limited to, those described above, by using a variety of available computer programs (*e.g.*, BLAST, from NCBI). Candidate homologs can be screened for biological activity using any suitable assay, including, but not limited to, those described herein.

2. Utrophin

In some embodiments, the spectrin-like repeats of the gene constructs of the present invention are replaced with spectrin-like repeats from another peptide (*e.g.*, homologs of dystrophin). For example, in some embodiments, spectrin-like repeats from the utrophin protein (*See e.g.*, Genbank accession number X69086; SEQ ID NO:3; Figure 3) are utilized. Utrophin is an autosomally-encoded homolog of dystrophin and has been postulated that the proteins play a similar physiological role (For a recent review, *See e.g.*, Blake *et al.*, Brain Pathology, 6:37 [1996]). Human utrophin shows substantial homology to dystrophin, with the major difference occurring in the rod domain, where utrophin lacks repeats 15 and 19 and two hinge regions (*See e.g.*, Love *et al.*, Nature 339:55 [1989]; Winder *et al.*, FEBS Lett., 369:27 [1995]). Utrophin thus contains 22 spectrin-like repeats and two hinge regions. A comparison of the rod domain of Utrophin and Dystrophin is shown in Figure 38.

In addition, in some embodiments, spectrin-like repeats from a homolog of utrophin are utilized. Homologs of utrophin have been identified in a variety of organisms, including mouse (Genbank accession number Y12229; SEQ ID NO:4; Figure 4) and rat (Genbank accession number AJ002967). The nucleic acid sequence of these or additional homologs can be compared to the nucleic acid sequence of human utrophin using any suitable methods, including, but not limited to, those described above. Candidate spectrin-like repeats from human utrophin or utrophin homologs can be screened for biological activity using any suitable assay, including, but not limited to, those described herein.

3. Alpha-actinin

In some embodiments, spectrin-like repeats from Dystrophin are replaced with spectrin-like repeats from alpha-actinin. The microfilament protein alpha-actinin exists as a dimer. The N-terminal regions of both polypeptides, arranged in antiparallel orientation, comprise the actin-binding regions, while the C-terminal, larger parts consist of four spectrin-like repeats that interact to form a rod-like structure (*See e.g.*, Winkler *et al.*, Eur. J. Biochem., 248:193 [1997]). In some embodiments, human alpha-actinin spectrin-like repeats are utilized (Genbank accession number M86406; SEQ ID NO:87; Figure 16). In other embodiments, alpha-actinin homologs from other organisms are utilized (*e.g.*, mouse (Genbank accession number AJ289242); Xenopus (Genbank accession number BE576799); and rat (Genbank accession number AF190909).

B. Variants

Still other embodiments of the present invention provide mutant or variant forms of spectrin-like repeats (*i.e.*, muteins). It is possible to modify the structure of a peptide having an activity of spectrin-like repeats for such purposes as enhancing therapeutic or prophylactic efficacy, or stability (*e.g.*, *ex vivo* shelf life, and/or resistance to proteolytic degradation *in vivo*). Such modified peptides provide additional peptides having a desired activity of the subject spectrin-like repeats as defined herein. A modified peptide can be produced in which the amino acid sequence has been altered, such as by amino acid substitution, deletion, or addition.

Moreover, as described above, variant forms (*e.g.*, mutants) of the subject spectrin-like repeats are also contemplated as finding use in the present invention. For example, it is contemplated that an isolated replacement of a leucine with an isoleucine or valine, an aspartate with a glutamate, a threonine with a serine, or a similar replacement of an amino acid with a structurally related amino acid (*i.e.*, conservative mutations) will not have a major effect on the biological activity of the resulting molecule. Accordingly, some embodiments of the present invention provide

variants of spectrin-like repeats containing conservative replacements. Conservative replacements are those that take place within a family of amino acids that are related in their side chains. Genetically encoded amino acids can be divided into four families: (1) acidic (aspartate, glutamate); (2) basic (lysine, arginine, histidine); (3) nonpolar (alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan); and (4) uncharged polar (glycine, asparagine, glutamine, cysteine, serine, threonine, tyrosine). Phenylalanine, tryptophan, and tyrosine are sometimes classified jointly as aromatic amino acids. In similar fashion, the amino acid repertoire can be grouped as (1) acidic (aspartate, glutamate); (2) basic (lysine, arginine histidine), (3) aliphatic (glycine, alanine, valine, leucine, isoleucine, serine, threonine), with serine and threonine optionally be grouped separately as aliphatic-hydroxyl; (4) aromatic (phenylalanine, tyrosine, tryptophan); (5) amide (asparagine, glutamine); and (6) sulfur -containing (cysteine and methionine) (*See e.g.*, Stryer (ed.), *Biochemistry*, 2nd ed, WH Freeman and Co. [1981]). Whether a change in the amino acid sequence of a peptide results in a functional homolog can be readily determined by assessing the ability of the variant peptide to function in a fashion similar to the wild-type protein. Peptides in which more than one replacement has taken place can readily be tested in the same manner.

The present invention further contemplates a method of generating sets of combinatorial mutants of the present spectrin-like repeats, as well as truncation mutants, and is especially useful for identifying potential variant sequences (*i.e.*, homologs) that possess the biological activity of spectrin-like repeats (*e.g.*, a decrease in muscle necrosis). In addition, screening such combinatorial libraries is used to generate, for example, novel spectrin-like repeat homologs that possess novel biological activities all together.

Therefore, in some embodiments of the present invention, spectrin-like repeat homologs are engineered by the present method to produce homologs with enhanced biological activity. In other embodiments of the present invention, combinatorially-derived homologs are generated which provide spectrin-like repeats that are easier to express and transfer to host cells. Such spectrin-like repeats, when

expressed from recombinant DNA constructs, can be used in therapeutic embodiments of the invention described below.

Still other embodiments of the present invention provide spectrin-like repeat homologs which have intracellular half-lives dramatically different than the corresponding wild-type protein. For example, the altered proteins comprising the spectrin-like repeat homologs are rendered either more stable or less stable to proteolytic degradation or other cellular process that result in destruction of, or otherwise inactivate spectrin-like repeats. Such homologs, and the genes that encode them, can be utilized to alter the pharmaceutical activity of constructs expressing spectrin-like repeats by modulating the half-life of the protein. For instance, a short half-life can give rise to more transient biological effects. As above, such proteins find use in pharmaceutical applications of the present invention.

In some embodiments of the combinatorial mutagenesis approach of the present invention, the amino acid sequences for a population of spectrin-like repeat homologs are aligned, preferably to promote the highest homology possible. Such a population of variants can include, for example, spectrin-like repeat homologs from one or more species, or spectrin-like repeat homologs from different proteins of the same species (*e.g.*, including, but not limited to, those described above). Amino acids that appear at each position of the aligned sequences are selected to create a degenerate set of combinatorial sequences.

In a preferred embodiment of the present invention, the combinatorial spectrin-like repeat library is produced by way of a degenerate library of genes encoding a library of polypeptides that each include at least a portion of candidate spectrin-like repeat sequences. For example, a mixture of synthetic oligonucleotides is enzymatically ligated into gene sequences such that the degenerate set of candidate spectrin-like repeat sequences are expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (*e.g.*, for phage display) containing the set of spectrin-like repeat sequences therein.

There are many ways by which the library of potential spectrin-like repeat homologs can be generated from a degenerate oligonucleotide sequence. In some

embodiments, chemical synthesis of a degenerate gene sequence is carried out in an automatic DNA synthesizer, and the synthetic genes are ligated into an appropriate gene for expression. The purpose of a degenerate set of genes is to provide, in one mixture, all of the sequences encoding the desired set of potential spectrin-like repeat sequences. The synthesis of degenerate oligonucleotides is well known in the art (*See e.g.*, Narang, Tetrahedron Lett., 39:3 9 [1983]; Itakura *et al.*, Recombinant DNA, *in* Walton (ed.), *Proceedings of the 3rd Cleveland Symposium on Macromolecules*, Elsevier, Amsterdam, pp 273-289 [1981]; Itakura *et al.*, Annu. Rev. Biochem., 53:323 [1984]; Itakura *et al.*, Science 198:1056 [1984]; Ike *et al.*, Nucl. Acid Res., 11:477 [1983]). Such techniques have been employed in the directed evolution of other proteins (*See e.g.*, Scott *et al.*, Science, 249:386-390 [1980]; Roberts *et al.*, Proc. Natl. Acad. Sci. USA, 89:2429-2433 [1992]; Devlin *et al.*, Science, 249: 404-406 [1990]; Cwirla *et al.*, Proc. Natl. Acad. Sci. USA, 87: 6378-6382 [1990]; as well as U.S. Pat. Nos. 5,223,409, 5,198,346, and 5,096,815, each of which is incorporated herein by reference). A wide range of techniques are known in the art for screening gene products of combinatorial libraries made by point mutations, and for screening cDNA libraries for gene products having a certain property. Such techniques are generally adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of spectrin-like repeat homologs. The most widely used techniques for screening large gene libraries typically comprise cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates relatively easy isolation of the vector encoding the gene whose product was detected. Each of the illustrative assays described below are amenable to high through-put analysis as necessary to screen large numbers of degenerate sequences created by combinatorial mutagenesis techniques.

Accordingly, in one embodiment of the present invention, the candidate genes comprising altered spectrin-like repeats are displayed on the surface of a cell or viral particle, and the ability of particular cells or viral particles to bind to a another

member of the DPC complex (*e.g.*, actin) is assayed. In other embodiments of the present invention, the gene library is cloned into the gene for a surface membrane protein of a bacterial cell, and the resulting fusion protein detected by panning (WO 88/06630; Fuchs *et al.*, BioTechnol., 9:1370 [1991]; and Goward *et al.*, TIBS 18:136 [1992]). In other embodiments of the present invention, fluorescently labeled molecules that bind proteins comprising spectrin like repeats (*e.g.*, actin), can be used to score for potentially functional spectrin-like repeat homologs. Cells are visually inspected and separated under a fluorescence microscope, or, where the morphology of the cell permits, separated by a fluorescence-activated cell sorter.

In an alternate embodiment of the present invention, the gene library is expressed as a fusion protein on the surface of a viral particle. For example, foreign peptide sequences are expressed on the surface of infectious phage in the filamentous phage system, thereby conferring two significant benefits. First, since these phage can be applied to affinity matrices at very high concentrations, a large number of phage can be screened at one time. Second, since each infectious phage displays the combinatorial gene product on its surface, if a particular phage is recovered from an affinity matrix in low yield, the phage can be amplified by another round of infection. The group of almost identical *E. coli* filamentous phages M13, fd, and fl are most often used in phage display libraries, as either of the phage gIII or gVIII coat proteins can be used to generate fusion proteins without disrupting the ultimate packaging of the viral particle (*See e.g.*, WO 90/02909; WO 92/09690; Marks *et al.*, J. Biol. Chem., 267:16007 [1992]; Griffths *et al.*, EMBO J., 12:725 [1993]; Clackson *et al.*, Nature, 352:624 [1991]; and Barbas *et al.*, Proc. Natl. Acad. Sci., 89:4457 [1992]).

In another embodiment of the present invention, the recombinant phage antibody system (*e.g.*, RPAS, Pharmacia Catalog number 27-9400-01) is modified for use in expressing and screening of spectrin-like repeat combinatorial libraries. The pCANTAB 5 phagemid of the RPAS kit contains the gene that encodes the phage gIII coat protein. In some embodiments of the present invention, the spectrin-like repeat combinatorial gene library is cloned into the phagemid adjacent to the gIII signal

sequence such that it is expressed as a gIII fusion protein. In other embodiments of the present invention, the phagemid is used to transform competent *E. coli* TG1 cells after ligation. In still other embodiments of the present invention, transformed cells are subsequently infected with M13KO7 helper phage to rescue the phagemid and its candidate spectrin-like repeat gene insert. The resulting recombinant phage contain phagemid DNA encoding a specific candidate spectrin-like repeat and display one or more copies of the corresponding fusion coat protein. In some embodiments of the present invention, the phage-displayed candidate proteins that are capable of, for example, binding to actin, are selected or enriched by panning. The bound phage is then isolated, and if the recombinant phage express at least one copy of the wild type gIII coat protein, they will retain their ability to infect *E. coli*. Thus, successive rounds of reinfection of *E. coli* and panning will greatly enrich for spectrin-like repeat homologs, which can then be screened for further biological activities.

In light of the present disclosure, other forms of mutagenesis generally applicable will be apparent to those skilled in the art in addition to the aforementioned rational mutagenesis based on conserved versus non-conserved residues. For example, spectrin-like repeat homologs can be generated and screened using, for example, alanine scanning mutagenesis and the like (Ruf *et al.*, Biochem., 33:1565 [1994]; Wang *et al.*, J. Biol. Chem., 269:3095 [1994]; Balint *et al.* Gene 137:109 [1993]; Grodberg *et al.*, Eur. J. Biochem., 218:597 [1993]; Nagashima *et al.*, J. Biol. Chem., 268:2888 [1993]; Lowman *et al.*, Biochem., 30:10832 [1991]; and Cunningham *et al.*, Science, 244:1081 [1989]), by linker scanning mutagenesis (Gustin *et al.*, Virol., 193:653 [1993]; Brown *et al.*, Mol. Cell. Biol., 12:2644 [1992]; McKnight *et al.*, Science, 232:316); or by saturation mutagenesis (Meyers *et al.*, Science, 232:613 [1986]).

C. Truncations and Additions

In yet other embodiments of the present invention, the spectrin-like repeats of human dystrophin are replaced by truncation or additions of spectrin-like repeats from

dystrophin or another protein, including, but not limited to, those described above. Accordingly, in some embodiments, amino acids are truncated from either end of one or more of the spectrin-like repeats in a given construct. The activity of truncation mutants is determined using any suitable assay, including, but not limited to, those disclosed herein.

In some embodiments, additional amino acids are added to either or both ends of the spectrin-like repeats in a given construct. In some embodiments, single amino acids are added and the activity of the construct is determined. Amino acids may be added to one or more of the spectrin-like repeats in a given construct. The activity of spectrin-like repeats comprising additional amino acids is determined using any suitable assay, including, but not limited to, those disclosed herein.

III. Carboxy-Terminal Domain Truncated Dystrophin Genes

In some embodiments, the present invention provides compositions comprising nucleic acid, wherein the nucleic acid encodes a mini-dystrophin peptide, and wherein the mini-dystrophin peptide comprises a substantially deleted dystrophin C-terminal domain (*e.g.*, 55% of the dystrophin C-terminal domain is missing). In some embodiments, this type of truncation prevents the mini-dystrophin peptide from binding both syntrophin and dystrobrevin.

The dystrophin COOH-terminal domain is located adjacent to the cysteine-rich domain, and contains an alternatively spliced region and two coiled-coil motifs (Blake *et al., Trends Biochem. Sci.*, 20:133, 1995). The alternatively spliced region binds three isoforms of syntrophin in muscle, while the coiled-coil motifs bind numerous members of the dystrobrevin family (Sadoulet-Puccio *et al., PNAS*, 94:12413, 1997). The dystrobrevins display significant homology with the COOH-terminal region of dystrophin, and the larger dystrobrevin isoforms also bind to the syntrophins. The importance and functional significance of syntrophin and dystrobrevin remains largely unknown, although they may be involved in cell signaling pathways (Grady *et al., Nat. Cell. Biol*, 1:215, 1999).

Researchers have previously generated transgenic *mdx* mouse strains expressing dystrophins deleted for either the syntrophin or the dystrobrevin binding domain (Rafael *et al., Hum. Mol. Genet.*, 3:1725, 1994; and Rafael *et al., J. Cell Biol.*, 134:93 1996). These mice displayed normal muscle function and essentially normal localization of syntrophin, dystrobrevin, and nNOS. Thus, while dystrobrevin appears to protect muscle from damage (Grady *et al., Nat. Cell. Biol*, 1:215, 1999), removal of the dystrobrevin binding site from dystrophin does not result in a dystrophy. Subsequent studies revealed that syntrophin and dystrobrevin bind each other in addition to dystrophin, so that removal of only one of the two binding sites on dystrophin might not sever the link between dystrophin, syntrophin and dystrobrevin. Surprisingly, the transgenic mice according to the present invention (*See* Example 1) displayed normal muscle function even though they lacked both the syntrophin and dystrobrevin binding sites.

IV. MCK Regulatory Regions

In certain embodiments, nucleic acid encoding mini-dystrophin peptides of the present invention are operably linked to muscle creatine kinase gene (MCK) regulatory regions and control elements, as well as mutated from of these regions and elements. (*see* U.S. Prov. App. Ser. No. 60/218,436, hereby incorporated by reference). In some embodiments, the nucleic acid encoding mini-dystrophin peptides is operably linked to these sequences to provide muscle specificity and reduced size such that the resulting construct is able to fit into, for example, a viral vector (*e.g.* adeno-associated virus). MCK gene regulatory regions (*e.g.* promoters and enhancers) display striated muscle-specific activity and have been characterized *in vitro* and *in vivo*. The major known regulatory regions in the mouse MCK gene include a 206 base pair muscle-specific enhancer located approximately 1.1 kb 5' of the transcription start site in mouse (*i.e.* SEQ ID NO:87) and a 358 base pair proximal promoter (*i.e.* SEQ ID NO:93) [Shield, *et al., Mol. Cell. Biol.*, 16:5058 (1996)]. A larger MCK promoter region may also be

employed (*e.g.* SEQ ID NO:92), as well as smaller MCK promoter regions (*e.g.* SEQ ID NO:94).

The 206 base pair MCK enhancer (SEQ ID NO:87) contains a number of sequence motifs, including two classes of E-boxes (MCK-L and MCK-R), CarG, and AT-rich sites. Similar E-box sequences are found in the enhancers of the human, rat, and rabbit MCK genes [*See*, Trask, *et al.*, *Nucleic Acids Res.*, 20:2313 (1992)]. Mutation may be made to this sequence by, for example, inserting an additional MCK-R control element into a wild-type enhancer sequence naturally containing one MCK-R control element (such that the resulting sequence has at least two MCK-R control elements). For example, the inserted MCK-R control element replaces the endogenous MCK-L control element. The 206 base pair mouse enhancer (SEQ ID NO:2) may be modified by replacing the left E-box (MCK-L) with a right E-Box (MCK-R) to generate a mutant muscle-specific enhancer region (*e.g.* to generate SEQ ID NO:88). A similar approximately 200 base pair wild type enhancer region in human may be modified by replacing the left E-box with a MCK-R to generate a mutant muscle-specific enhancer region (*e.g.* 2R human enhancer regions).

Another modification that may be made to generate mutant muscle-specific enhancer regions by inserting the S5 sequence GAGCGGTTA (SEQ ID NO:95) into wild type mouse, human, and rat enhancer sequence. Making such a modification to the mouse enhancer SEQ ID NO:87, for example, generates S5 mutant muscle-specific enhancer regions (*e.g.* SEQ ID NO:89). Another modification that may be made, for example, to the wild type mouse enhancer is replacing the left E-box (MCK-L) with a right E-Box (MCK-R), and also inserting the 5S sequence, to generate 2R5S type sequences (*e.g.* in mouse, SEQ ID NO:90). These mutant muscle-specific enhancer regions may have additional sequences added to them or sequences that are taken away. For example, the mutant muscle-specific enhancer regions may have a portion of the sequence removed (*e.g.* the 3' 41 base pairs). Examples of such mutant truncation 2RS5 sequences in mouse is SEQ ID NO:91 with the 3' 41 base pairs removed, generating mutant truncated 2RS5 muscle-specific enhancer regions.

Any of these wild-type or mutant muscle-specific enhancer regions described above may be further modified to produce additional mutants. These additional mutants include, but are not limited to, muscle-specific enhancer regions having deletions, insertions or substitutions of different nucleotides or nucleotide analogs so long as the transcriptional activity of the enhancer region is maintained. Guidance in determining which and how many nucleotide bases may be substituted, inserted or deleted without abolishing the transcriptional activity may be found using computer programs well known in the art, for example, DNAStar software or GCG (Univ. of Wisconsin) or may be determined empirically using assays provided by the present invention.

V. Expression Vectors

The present invention contemplates the use of expression vectors with the compositions and methods of the present invention (e.g. with the nucleic acid constructs encoding the mini-dystrophin peptides). Vectors suitable for use with the methods and compositions of the present invention, for example, should be able to adequately package and carry the compositions and cassettes described herein. A number of suitable vectors are known in the art including, but are not limited to, the following: 1) Adenoviral Vectors; 2) Second Generation Adenoviral Vectors; 3) Gutted Adenoviral Vectors; 4) Adeno-Associated Virus Vectors; and 5) Lentiviral Vectors.

Those skilled in the art will recognize and appreciate that other vectors are suitable for use with methods and compositions of the present invention. Indeed, the present invention is not intended to be limited to the use of the recited vectors, as such, alternative means for delivering the compositions of the present invention are contemplated. For example, in various embodiments, the compositions of the present invention are associated with retrovirus vectors and herpes virus vectors, plasmids, cosmids, artificial yeast chromosomes, mechanical, electrical, and chemical transfection methods, and the like. Exemplary delivery approaches are discussed below.

1. **Adenoviral Vectors**

Self-propagating adenovirus (Ad) vectors have been extensively utilized to deliver foreign genes to a great variety of cell types *in vitro* and *in vivo*. "Self-propagating viruses" are those which can be produced by transfection of a single piece of DNA (the recombinant viral genome) into a single packaging cell line to produce infectious virus; self-propagating viruses do not require the use of helper virus for propagation. As with many vectors, adenoviral vectors have limitations on the amount of heterologous nucleic acid they are capable of delivering to cells. For example, the capacity of adenovirus is approximately 8-10 kb, the capacity of adeno-associated virus is approximately 4.8 kb, and the capacity of lentivirus is approximately 8.9 kb. Thus, the mutants of the present invention that provide shorter nucleic acid sequences encoding the mini-dystrophin peptides (compared to full length wild-type dystrophin (14kb)), improve the carrying capacity of such vectors.

2. **Second Generation Adenoviral Vectors**

In an effort to address the viral replication problems associated with first generation Ad vectors, so called "second generation" Ad vectors have been developed. Second generation Ad vectors delete the early regions of the Ad genome (E2A, E2B, and E4). Highly modified second generation Ad vectors are less likely to generate replication-competent virus during large-scale vector preparation, and complete inhabitation of Ad genome replication should abolish late gene replication. Host immune response against late viral proteins is thus reduced [*See* Amalfitano *et al.*, "Production and Characterization of Improved Adenovirus Vectors With the E1, E2b, and E3 Genes Deleted," J. Virol. 72:926-933 (1998)]. The elimination of E2A, E2B, and E4 genes from the Ad genome also provide increased cloning capacity. The deletion of two or more of these genes from the Ad genome allows for example, the delivery of full length or cDNA dystrophin genes via Ad vectors [Kumar-Singh *et al, Hum. Mol. Genet.*, 5:913 (1996)].

3. Gutted Adenoviral Vectors

"Gutted," or helper dependent, Ad vectors contain *cis*-acting DNA sequences that direct adenoviral replication and packaging but do not contain viral coding sequences [*See* Fisher *et al.* "Recombinant Adenovirus Deleted of All Viral Genes for Gene Therapy of Cystic Fibrosis," *Virology* 217:11-22 (1996) and Kochanek *et al.* "A New Adenoviral Vector: Replacement of All Viral Coding Sequences With 28 kb of DNA Independently Expressing Both Full-length Dystrophin and Beta-galactosidase'" *Proc. Nat. Acad. Sci. USA* 93:5731-5736 (1996)]. Gutted vectors are defective viruses produced by replication in the presence of a helper virus, which provides all of the necessary viral proteins *in trans*. Since gutted vectors do not contain any viral genes, expression of viral proteins is not possible.

Recent developments have advanced the field of gutted vector production [*See* Hardy *et al.*, "Construction of Adenovirus Vectors Through Cre-lox Recombination," J. Virol. 71:1842-1849 (1997) and Hartigan-O'Conner *et al.*, "Improved Production of Gutted Adenovirus in Cells Expressing Adenovirus Preterminal Protein and DNA Polymerase," *J. Virol.* 73:7835-7841 (1999)]. Gutted Ad vectors are able to maximally accommodate up to about 37 kb of exogenous DNA, however, 28-30 kb is more typical. For example, a gutted Ad vector can accommodate the full length dystrophin or cDNA, but also expression cassettes or modulator proteins.

4. Adeno-Associated Virus Vectors

In preferred embodiments, the nucleic acid encoding the mini-dystrophin peptides of the present invention are inserted in adeno-associated vectors (AAV vectors). AAV vectors evade a host's immune response and achieve persistent gene expression through avoidance of the antigenic presentation by the host's professional APCs such as dendritic cells. Most AAV genomes in muscle tissue are present in the form of large circular multimers. AAV's are only able to carry about 5 kb of exogenous DNA. As such, the nucleic acid of the present invention encoding the

mini-dystrophin peptides is well suited, in some embodiments, for insertion into these vectors due the reduced size of the nucleic acid sequences.

The dystrophin expression cassettes of the present invention (containing nucleic acid encoding mini-dystrophin peptides) may be cloned into any of a variety of cis-acting plasmid vectors that contain the adeno-associated virus inverted terminal repeats (ITRs) to allow production of infectious virus. For example, one such plasmid is the cis-acting plasmid (pCisAV) (Yan *et al., PNAS*, 97:6716-6721, 2000). This plasmid contains the AAV-ITRs separated by a NotI cloning site. The ITR elements were derived from pSub201, a recombinant plasmid from which an infectious adeno-associated virus genome can be excised *in vitro* and used to study viral replication. After ligation of the dystrophin expression cassette (isolated as a NotI fragment from pCK6DysR4-23-71-78An) into NotI-digested pCisAV, rAAV stocks are generated by cotransfection of pCisAV. CK6DysR4-23-71-78An and pRep/Cap (Fisher, *et al., J. Virol.* 70:520-532, 1996) together with coinfection of the recombinant adenovirus Ad.CMVlacZ into 293 cells. Recombinant AAV vector, for example, may then be purified on CsCl gradients as described (Duan, *et al., Virus Res.* 48:41-56, 1997).

5. **Lentiviral Vectors**

Vectors based on human or feline lentiviruses have emerged as another vector useful for gene therapy applications. Lentivirus-based vectors infect nondividing cells as part of their normal life cycles, and are produced by expression of a package-able vector construct in a cell line that expresses viral proteins. The small size of lentiviral particles constrains the amount of exogenous DNA they are able to carry to about 10 kb. However, once again, the small size nucleic acid encoding the mini-dystrophin peptides of the present invention allow such vectors to be employed.

6. **Retroviruses**

Vectors based on Moloney murine leukemia viruses (MMLV) and other retroviruses have emerged as useful for gene therapy applications. These vectors

stably transduce actively dividing cells as part of their normal life cycles, and integrate into host cell chromosomes. Retroviruses may be employed with the compositions of the present invention (e.g. gene therapy), for example, in the context of infection and transduction of muscle precursor cells such as myoblasts, satellite cells, or other muscle stem cells.

EXPERIMENTAL

The following examples are provided in order to demonstrate and further illustrate certain preferred embodiments and aspects of the present invention and are not to be construed as limiting the scope thereof.

In the experimental disclosure which follows, the following abbreviations apply: N (normal); M (molar); mM (millimolar); μM (micromolar); mol (moles); mmol (millimoles); μmol (micromoles); nmol (nanomoles); pmol (picomoles); g (grams); mg (milligrams); μg (micrograms); ng (nanograms); l or L (liters); ml (milliliters); μl (microliters); cm (centimeters); mm (millimeters); μm (micrometers); nm (nanometers); °C (degrees Centigrade); Sigma (Sigma Chemical Co., St. Louis, MO); and

EXAMPLE 1

Carboxy-Terminal Domain Truncated Dystrophin Genes

This example describes the generation of carboxy-terminal truncated dystrophin nucleic acid sequences. In particular, this examples describes the construction of dystrophin nucleic acid sequence with the entire carboxy-terminal domain deleted, and testing of this sequence in a mouse model for DMD.

A. Methods

The bases encoding amino acids 3402–3675 (corresponding to exons 71–78) were deleted from the full length murine dystrophin cDNA (SEQ ID NO:2, accession

No. M68859) by recombinant PCR, leaving the last three amino acids (exon 79) of the dystrophin protein unaltered. This dystrophin Δ71–78 cDNA was cloned into an expression vector containing bases -2139 to +239 of the human -skeletal actin (HSA) promoter (Brennan, *et al., J. Biol. Chem.* 268:719, 1993). A splice acceptor from the SV40 VP1 intron (isolated as a 400 bp HindIII/XbaI fragment from pSVL; Amersham Pharmacia Biotech) was inserted immediately 3' of the HSA fragment, and the SV40 polyadenylation signal (isolated as a BamHI fragment from pCMVß; MacGregor and Caskey, Nuc. Acid. Res., 17:2365, 1989) was inserted 3' of the dystrophin cDNA. The excised dystrophin Δ71–78 expression cassette was injected into wild-type C57Bl/10 x SJL/J F2 hybrid embryos, and F_0 mice were screened by PCR. Five positive F_0's were backcrossed onto the C57Bl/10mdx background, and the line with the most uniform expression levels was selected for analysis. Also employed were previously described transgenic *mdx* mice that express dystrophin constructs deleted approximately for exons 71–74 (Δ71–74) or exons 75–78 (Δ75–78), which remove amino acids 3402–3511 and 3528–3675, respectively, *See* Rafael *et al., J. Cell Biol.*, 134:93-102, 1996). Transgenic *mdx* line Dp71 expresses the Dp71 isoform of dystrophin in striated muscle (Cox *et al., Nat. Genet.*, 8:333-339, 1994).

i. Morphology Methods

Quadriceps, soleus, extensor digitorum longus (EDL), tibialis anterior, and diaphragm muscles were removed from the mice, frozen in liquid nitrogen cooled O.C.T. embedding medium (Tissue-Tek), and cut into 7-µm sections. After fixing in 3.7% formaldehyde, sections were stained in hematoxylin and eosin–phloxine. Stained sections were imaged with a Nikon E1000 microscope connected to a Spot-2 CCD camera. To determine the percentage of fibers containing central nuclei, the number of muscle fibers with centrally-located nuclei was divided by the total number of muscle fibers.

ii. Evans Blue Assays

4 month old control mice and 71–78 mice were analyzed after injection with Evans blue, as described previously (Straub *et al., J. Cell. Biol.*, 139:375-385, 1997). In brief, mice were tail vein–injected with 150 µl of a solution containing 10 mg/ml Evans blue dye in PBS (150 mM NaCl, 50 mM Tris, pH 7.4). After 3 hours, the animals were euthanized and mouse tissues were either fixed in 3.7% formaldehyde/0.5% glutaraldehyde to observe gross dye uptake, or frozen unfixed in O.C.T. embedding medium. To examine Evans blue uptake by individual fibers, 7-µm-thick frozen sections were fixed in cold acetone and analyzed by fluorescence microscopy.

iii. Immunofluorescence Assays

Quadriceps and diaphragm muscles from C57Bl/10, *mdx*, and Δ71–78 mice were removed, frozen in O.C.T. embedding medium, and cut into 7-µm sections. Immunofluorescence was performed with previously described antibodies against dystrophin (NH_2 terminus), α1-syntrophin (SYN17), ß1-syntrophin, α-dystrobrevin-1 (DB670), α-dystrobrevin-2 (DB2), and utrophin. After incubation with primary antibodies, cryosections were incubated with an FITC-conjugated goat anti–rabbit secondary antibody and fluorescent images were viewed on a Nikon E1000 microscope. Antibodies to α-sarcoglycan (Rabbit 98), ß-sarcoglycan (Goat 26), γ-sarcoglycan (Rabbit 245), δ-sarcoglycan (Rabbit 215), sarcospan (Rabbit 235), α-dystroglycan (Goat 20), ß-dystroglycan (AP 83), or nNOS (Rabbit 200) have been described previously (Duclos *et al., J. Cell. Biol.*, 142:1461, 1998). Cy3-conjugated secondary antibodies were used and images were viewed on a Bio-Rad MRC-600 laser scanning confocal microscope. All digitized images were captured under the same conditions.

iv. Measurements of Contractile Properties Methods

Contractile properties of muscles from 6-month-old Δ71–78 transgenic mice were compared with those of C57Bl/10 wild-type and *mdx* mice using methods

described previously (Lynch *et al., Am. J. Physiol.*, 272:C2063, 1997). The samples included eight muscles each from the EDL, soleus, and diaphragm. Mice were deeply anesthetized with avertin and each muscle was isolated and dissected free from the mouse. After removal of the limb muscles, the mice were euthanized with the removal of the diaphragm muscle. The muscles were immersed in a bath filled with oxygenated buffered mammalian Ringer's solution (137 mM NaCl, 24 mM $NaHCO_3$, 11 mM glucose, 5 mM KCl, 2 mM $CaCl_2$, 1 mM $MgSO_4$, 1 mM NaH_2PO_4, and 0.025 mM tubocurarine chloride, pH 7.4). For each muscle, one tendon was tied to a servomotor and the other tendon to a force transducer. Muscles were stretched from slack length to the optimal length for force development and then stimulated at a frequency that produced absolute isometric tetanic force (mN). After the measurements of the contractile properties, the muscles were removed from the bath, blotted and weighed to determine muscle mass. Specific force (kN/m^2) was calculated by dividing absolute force by total fiber cross sectional area.

v. Muscle Membrane Isolation Methods

Muscle microsomes from 12–14 month-old C57Bl/10, *mdx*, Δ71–78, Δ71–74, Δ75–78, and Dp71 mice were prepared as described previously (Ohlendieck *et al., J. Cell. Biol.*, 112:135, 1991). In brief, skeletal muscle was homogenized in 7.5-vol homogenization buffer plus protease inhibitor Complete (Boehringer). The homogenate was centrifuged at 14,000 g for 15 min to remove cellular debris. The supernatant was filtered through cheesecloth and spun at 142,000 g for 37 minutes to collect microsomes. The microsome pellet was resuspended in KCl wash buffer (0.6 M KCl, 0.3 M sucrose, 50 mM Tris-HCl, pH 7.4) plus protease inhibitors and recentrifuged at 142,000 g for 37 minutes to obtain KCl-washed microsomes. The final pellet was resuspended in 0.3 M sucrose and 20 mM Tris-maleate, pH 7.0. Samples were quantified by the Coomassie Plus Protein Assay Reagent (Pierce Chemical Co.) and equivalent protein loading was verified by SDS-PAGE. KCl-washed microsomes were analyzed by Western blot using antibodies against

ß2-syntrophin, pan syntrophin, nNOS (Transduction Laboratories), ß-dystroglycan, α-sarcoglycan (Novocastra Laboratories), and other proteins described above.

B. Results

i. Generation of Dystrophin Δ71–78 Transgenic Mice

To test the function of a dystrophin protein lacking both the syntrophin and dystrobrevin binding sites, we prepared a cDNA expression vector deleted for the COOH-terminal domain (corresponding to exons 71–78; *See* Fig 19) as described above. The structure of several dystrophin transgenic constructs previously tested are also shown for comparison. Mice expressing the dystrophin Δ71–78 transgene were crossed onto the *mdx* background and dystrophin levels were analyzed by Western blotting. The expression of the dystrophin Δ71–78 transgene in skeletal muscle was determined to be 10-fold higher than endogenous dystrophin. Immunofluorescent staining of quadriceps muscle using an antibody against the NH_2-terminus of dystrophin revealed that the Δ71–78 protein was localized to the sarcolemma, similar to wild-type dystrophin. Dystrophin Δ71–78 expression was also found to be uniform in the diaphragm, EDL, and soleus muscles, but the tibialis anterior muscle displayed a mosaic expression pattern. The human skeletal muscle -actin promoter used in this study was not expressed in either smooth or cardiac muscle.

ii. Morphology of Dystrophin Δ71–78 Mice Appears Normal

We initially analyzed transgenic *mdx* mouse muscle tissues for morphological signs of dystrophy. Hematoxylin and eosin–stained limb and diaphragm skeletal muscle sections of dystrophin Δ71–78 mice revealed none of the signs of fibrosis, necrotic fibers, or mononuclear cell infiltration that were apparent in age-matched mdx controls. NMJs (neuromuscular junctions) of transgenic mice stained with rhodamine-labeled -bungarotoxin consistently appeared normal in contrast to the varying degrees of postsynaptic folding observed in *mdx* NMJs. *Mdx* muscle fibers have previously been shown to be highly permeable to the vital dye Evans blue *in*

vivo, reflecting damage to the dystrophic fiber sarcolemma (Matsuda *et al., J. Biochem.* (Tokyo), 118:959, 1995). Skeletal muscle fibers from dystrophin Δ71–78 mice, like wild-type animals, were found not to be permeable to Evans blue dye.

iii. Analysis of Centrally Nucleated Muscle Fibers

Another hallmark of dystrophy in *mdx* mice is the presence of large numbers of centrally-nucleated muscle fibers, reflecting cycles of fiber degeneration and regeneration (Torres and Duchen, *Brain*, 110:269, 1987). To estimate the degree of myofiber regeneration occurring in Δ71–78 transgenic mice, centrally nucleated fibers were counted from a variety of muscle groups in age-matched wild-type, *mdx*, and Δ71–78 mice (*See*, Table 2). By 4 months of age, 71% of muscle fibers in *mdx* quadriceps muscles contained central nuclei, whereas wild-type muscles had <1%. Interestingly, 4 month old dystrophin Δ71–78 quadriceps muscles displayed 1% central nuclei, indicating that very little, if any, regeneration was occurring. When 1-year-old mice were compared, a modest increase in centrally nucleated fibers became apparent. Quadriceps muscles from Δ71–78 mice contained 10% centrally nucleated fibers, although diaphragm muscles still displayed <1%. EDL and soleus muscles displayed 5 and 8% centrally nucleated fibers, respectively. For comparison, 1-year-old wild-type mice had <1% centrally nucleated fibers in both limb and diaphragm muscles. Furthermore, 1-year-old *mdx* limb muscles had 60% centrally nucleated fibers, whereas the diaphragm had 35%.

Table 2

Percentage of Centrally Nucleated Fibers in Mouse Skeletal Muscles

Line	Age	Quad	Dia	TA	EDL	Soleus
C57/B110	4	<1	<1	ND	ND	ND
mdx	4	71	58	ND	ND	ND
Δ71-78	4	1	<1	ND	ND	ND
C57/B110	12	<1	<1	<1	<1	<1
mdx	12	65	35	58	50	61
Δ71-78	12	10	<1	ND	5	8
Δ71-74	15	5	<1	<1	<1	ND
Δ75-78	15	8	<1	4	2	7

Quad = quadriceps; Dia = diaphragm; TA = tibialis anterior; Age is in months

Previous studies of transgenic mice expressing dystrophins deleted for exons Δ71–74 (Δ71–74) or exons Δ75–78 (Δ75–78) revealed no increase in the numbers of centrally nucleated fibers by 4 months of age (Rafael *et al.* 1996, *see* above). To contrast these mice with the 71–78 transgenics, central nuclei counts were performed on 15-month-old Δ71–74 and 75–78 mice. It was determined that these animals had central nuclei counts in between those of wild-type and Δ71–78 mice. The Δ71–74 and Δ75–78 mice had 5 and 8% centrally nucleated fibers in quadriceps, respectively (Table 2).

iv. Contractile Properties

Compared with muscles of wild-type mice, those from *mdx* mice displayed a significant amount of necrosis, fibrosis, and infiltrating mononuclear cells. *mdx* skeletal muscles also displayed a loss of specific force–generating capacities when muscles were stimulated to contract *in vitro*, providing an extremely sensitive and quantitative measurement of the dystrophic process (Fig 20 A). In contrast, dystrophin Δ71–78 mice had no major abnormalities when subjected to the same analysis (Fig 20

B). Muscle mass for both EDL and diaphragm were not significantly different between dystrophin Δ71–78 and wild-type mice, whereas dystrophin Δ71–78 soleus muscles were slightly hypertrophied. When stimulated to contract, all three muscle groups displayed specific forces not significantly different from wild-type ($P < 0.05$). These results demonstrate that the dystrophin Δ71–78 protein has essentially the same functional capacity as the full-length protein.

v. Localization of the DAP Complex in Δ71–78 Mice

Immunofluorescent analysis of the peripheral DAP complex revealed α1-syntrophin, ß1-syntrophin, α-dystrobrevin-1, and α-dystrobrevin-2 to be localized at the sarcolemma with dystrophin, despite the lack of syntrophin and dystrobrevin binding sites in the transgene-encoded dystrophin. α1-syntrophin levels were similar between wild-type and Δ71–78 mice. However, the levels of ß1-syntrophin were elevated at the membrane in Δ71–78 mice, particularly in those fibers that normally express significant levels of this isoform. α-dystrobrevin-1 was primarily located at the NMJ in wild-type mice, and was exclusively located at the NMJs in *mdx* mice. Surprisingly, in dystrophin Δ71–78 mice, higher levels of α-dystrobrevin-1 were observed at the sarcolemma than in wild-type mice. The Δ71–78 mice also displayed a slight increase in utrophin localization along the sarcolemma, but this increase was less than the increase in *mdx* fibers. Immunofluorescent localization of the sarcoglycans, α- and ß-dystroglycan, sarcospan, and nNOS in Δ71–78 mice revealed no differences in the expression of these proteins when compared with wild-type mice. The proper localization of these proteins to the sarcolemma indicated that membrane targeting of the DAP complex components can proceed in the absence of the COOH-terminal domain of dystrophin.

vi. DAP Complex Protein Levels

To examine the levels of the DAP complex members that associate with dystrophin, muscle microsomes were prepared from wild-type and dystrophin Δ71–78 mice and analyzed by Western blotting. This approach provides information on the

relative abundance of individual DAP complex members in muscles of separate lines of mice. Slightly elevated levels of ß-dystroglycan were detected in dystrophin Δ71–78 mice, which we have previously observed whenever dystrophin is overexpressed. Isoforms of syntrophin and dystrobrevin were present at slightly different levels when the dystrophin Δ71–78 membranes were compared with those from wild-type mice. α1-syntrophin and ß2-syntrophin levels were lower than in wild-type mice, whereas the level of ß1-syntrophin was elevated. Although there was approximately the same amount of α-dystrobrevin-2, there were elevated levels of α-dystrobrevin-1 in Δ71–78 microsomes. A reduction in nNOS was observed in dystrophin Δ71–78 muscle, indicating that nNOS binds weakly to the DAP complex in Δ71–78 mice. Levels of α-sarcoglycan were similar in all lines tested, and provided an internal control for protein loading.

Since some DAP complex members exhibited isoform changes in Δ71–78 mice, we examined purified microsomes from dystrophin Δ71–74 and Δ75–78 mice. Transgenic *mdx* mice that express the dystrophin isoform Dp71 in muscle were also included in this study since these dystrophic mice have the DAP complex present at the sarcolemma. α1-syntrophin levels were lower in all four transgenic lines compared with wild-type mice. Surprisingly, ß1-syntrophin was absent in Δ71–74 microsomes but was highly overexpressed in Δ75–78 and Dp71 microsomes. The Δ71–74 microsomes had equivalent ß2-syntrophin levels when compared with wild-type microsomes, but this isoform of syntrophin was reduced in both Δ75–78 and Dp71 microsomes. A pan syntrophin antibody, which detects all three isoforms of syntrophin, confirmed the upregulation of syntrophin in Δ75–78 and Dp71 microsomes. Similar to Δ71–78, α-dystrobrevin-1 was elevated in all dystrophin transgenic microsome preparations. However, in comparison with wild-type, α-dystrobrevin-2 was higher in Δ71–74 and Δ75–78, but equal in Dp71 microsomes. Contrary to the Δ71–78 mice, deleting either exons 71–74 or 75–78 restored nNOS to wild-type levels. However, Dp71 mice, which lack the NH_2-terminal and rod domains of dystrophin, did not retain nNOS in the microsome fractions. Previous studies have also shown that utrophin is upregulated in mdx and Dp71 mice (Ohlendieck et al.,

Neuron, 7:499-508, 1991). Therefore, utrophin levels were compared in all transgenic lines and we found that Δ71–78, Δ71–74, and Δ75–78 mice do not have the elevated levels seen in *mdx* and Dp71 mice.

EXAMPLE 2

Construction of ΔR4-R23, ΔR2-R21+H3, and ΔR2-R1

This example describes the construction of R4-R23 (micro-dys1), ΔR2-R21+H3 (micro-dys3), and ΔR2-R1 (micro-dys2), three sequences with 4 spectrin-like repeat encoding sequences. The 'full-length' human dystrophin cDNA that was started with was actually a sequence slightly smaller than the true full-length human dystrophin cDNA. In particular, the starting sequence, called full-length HDMD (SEQ ID NO:47, see Fig. 23) is the same as the wild-type human dystrophin in SEQ ID NO:1, except the 3' 1861 base pairs are deleted (at an XbaI site), and the 39 base pair alternatively spliced exon 71 (bases 10432-10470) are deleted. This sequence (SEQ ID NO:47) is originally in pBSX (SEQ ID NO:46, See Figs. 21 and 22).

A. Cloning ΔR4-R23

The procedure used for cloning ΔR4-R23 is outlined in Figure 24. Initially, three PCR reactions were performed (employing Pfu polymerase) to create the deletion shown in Figure 24. The primers employed in the first reaction were 5' GAA CAA GAT TCA CAC AAC TGG C 3' (SEQ ID NO:48), which anneals to 1954-1975 of the HDMD clone, and 5' GTT CCT GGA GTC TTT CAA **GAT CCA CAG TAA TCT GCC TC** 3' (SEQ ID NO:49), which is a reversed, tailed primer (the bold sequence anneals to 2359-2341 of the HDMD clone, and the underlined sequence anneals to 9023-9005 the HDMD clone. PCR was conducted employing these primers, and a 425 bp PCR product was produced. The first primer employed in the second reaction was **5' GAG GCA GAT TAC TGT GGA T**CT TGA AAG ACT CCA GGA AC 3' (SEQ ID NO:50), which is the reverse complement primer of SEQ ID NO:49 (the bold-faced sequence of SEQ ID NO:50) anneals to 2341-2359 of the HDMD clone in the forward direction. The underlined sequence anneals to 9005-9023 of the HDMD clone in the

forward direction. The other primer employed for the second reaction was 5' TGT TTG GCG AGA TGG CTC 3' (SEQ ID NO:51) which anneals to 9413-9396 of HDMD in the reverse direction. PCR was conducted employing these primers, and a 427 bp PCR product was produced. The third reaction employed the products from steps 1 and 2 and the outside primers SEQ ID NO:48 and SEQ ID NO:51, producing a 814 bp fragment by PCR. This fragment was then digested with NcoI and HindIII to produce a 581 bp DNA fragment.

This 581 bp fragment was then cloned into a 5016 bp NcoI + Hind III fragment from the HDMD clone. The 581 bp fragment contained part of repeat 3, all of Hinge 2, and part of repeat 24. The NcoI site used in the HDMD clone was located at 2055 bp. The Hind III site was located at 9281 bp. The 5016 fragment contained the pBSX cloning vector sequence, and the entire 5' UTR, the entire N terminus, Hinge 1, Repeats 1, 2, and part of repeat 3 up to the NcoI site of human dystrophin. Ligation of the 5016 bp fragment and 581 bp fragment (step 2) was then performed to created a 5597 bp sequence.

Step 3 was then performed to clone a 2.9 kb HindIII fragment containing part of repeat 24, the C terminus, and the 3' UTR (See Fig. 24). The 5' HindIII site is located at 9281 bp of the HDMD clone. The 3' HindIII site of this fragment is derived from pBSX polylinker. This 2.9 kb fragment was cloned into the HindIII site of the product of Step 2 to yield an 8.5 kb plasmid, composed of the ΔR4-R23 cDNA plus pBSX. The entire ΔR4-R23 cDNA was excised from pBSX with NotI and cloned into the NotI site of the HSA expression vector (HSA promoter - VP1 intron - NotI site - tandem SV40 poly adenylation site).

B. Cloning ΔR2-R21+H3

The procedure used for cloning ΔR2-R21+H3 is outlined in Figure 25. Initially, four PCR reactions were performed (employing Pfu polymerase) to create the deletion shown in Figure 25. The primers employed in the first reaction were 5' GAT GTG GAA GTG GTG AAA GAC 3 (SEQ ID NO:52), which anneals to 1319-1330 of the HDMD clone, and 5' CCA ATA GTG GTC AGT CCA GGA GCA TGT AAA

TTG CTT TG 3' (SEQ ID NO:53), which is a reverse, tailed primer (the bold-faced sequence anneals to 1546-1532 of the HDMD clone and the underlined sequence anneals to 7512-7490 of the HDMD clone. PCR was conducted with these primers and a 228 bp PCR product was produced. The first primer employed in the second reaction was 5' **CAA AGC AAT TTA CAT** GCT CCT GGA CTG ACC ACT ATT GG 3' (SEQ ID NO:54), which is the reverse complement of SEQ ID NO:53 (the bold-faced sequence of SEQ ID NO:54 anneals to 1532-1546 of the HDMD clone in the forward direction, and the underlined sequence anneals to 7512-7490 of the HDMD clone in the forward direction. The other primer employed in the second reaction was 5' **CTG TTG CAG TAA TCT ATG** CTC CAA CAT CAA GGA AGA TG 3' (SEQ ID NO:55), and the bold-faced sequence anneals to 8287-8270 of the HDMD clone, and the underlined sequence anneals to 7612-7593 of the HDMD clone as a reverse primer. PCR was performed with these primers, and a 123 bp PCR product was produced. The first primer employed in the third reaction was 5' CAT CTT CCT TGA TGT TGG **AGC ATA GAT TAC TGC AAC AG** 3' (SEQ ID NO:56), the underlined sequence anneals to 7593-7612 of the HDMD clone in the forward direction, and the bold-faced sequence anneals to 8270-8287. The second primer employed in the third reaction was SEQ ID NO:51 (see above), which anneals to 9413-9396 in the reverse direction. PCR was performed with these primers, and a 1143 bp fragment was produced. The fourth reaction employed the products from reactions 1,2, and 3 as template, and the outside primers (SEQ ID NO:52 and SEQ ID NO:51), and a 1494 bp fragment was produced using Pfu polymerase.

This 1494 bp fragment was then digested with MunI and HindIII to produce a 1270 bp band and cloned into a 4320 bp MunI + HindIII fragment from the HDMD clone. The 1270 bp fragment contained the part of repeat 1, all of hinge 3, repeat 22, repeat 23, and part of repeat 24. The 4320 bp fragment contained the 5' UTR of HDMD, the N terminus, Hinge 1, and part of repeat 1 and pBSX. The MunI site in HDMD is located at base 1409. The HindIII site is at 9281 bp. Ligation of the 4320 bp fragment and the 1270 bp fragment was then performed (See Figure 25) and a

4490 bp fragment was produced. Step 3 was performed as describe above for ΔR4-R23 to generate ΔR2-R21+H3.

C. Cloning ΔR2-R21

The cloning of ΔR2-R21 was performed essentially the same way as for ΔR2-R21+H3, with the exception of the recombinant PCR reaction to assemble the rod domain deletion (See, Figure 26). All other steps are the same. Three PCR reactions were performed (using Pfu polymerase) to create the deletion. The primers employed in the first reaction were SEQ ID NO:52 (see above), and 5' CTG TTG CAG TAA TCT ATG **ATG TAA ATT GCT TTG** 3' (SEQ ID NO:57), the underlined sequence anneals to 8287-8270 of the HDMD clone in the reverse direction, and the bold-faced sequence anneals to 1546-1532 of the HDMD clone in the reverse direction. PCR was performed with these primers, and a 250 bp product was obtained. The first primer employed in the second reaction was 5' **CAA AGC AAT TTA CAT** CAT AGA TTA CTG CAA CAG 3' (SEQ ID NO:58), which is is the reverse complement of SEQ ID NO:57 (the bold-faced sequence of SEQ ID NO:58 anneals to 1532-1546 of the HDMD clone in the forward direction, and the underlined sequence anneals to 8270-8287 of the HDMD clone in the forward direction. The other primer employed in the second reaction was SEQ ID NO:51, which anneals to 9413-9396 in the reverse direction. PCR was performed with these primers and a 1143 bp product was obtained. The third reaction employed the products from reactions 1 and 2 (as template) and the outside primers (SEQ ID NO:52 and SEQ ID NO:51), and a 1383 bp fragment was produced. This fragment was then digested with MunI and HindIII to produce an 1147 bp fragment containing part of repeat 1, repeat 22, repeat 23, and part of repeat 24. This was then cloned into the same MunI + HindIII HDMD fragment described for the ΔR2-R21+H3 clone and all other steps thereafter were the same.

EXAMPLE 3

ΔR4-R23 Deletions

This example describes the construction of 5' UTR, 3' UTR, and C-terminal deletions of ΔR4-R23 (making it even smaller), as well as the addition of polyadenylation and promoter sequences. This example also describes the alteration of the Kozak sequence (to become more like that of consensus).

A. Deletion of the 3' UTR

In order to delete the 3' UTR, the following two primers were employed 5' TCT CTC CAA GAT CAC CTC G 3' (SEQ ID NO:64), which anneals to 9117-9134 of the HDMD full length clone, and 5' ATG AAG CTT **GCG GCC GCA** TGC GGG AAT CAG GAG TTG 3' (SEQ ID NO:65) (the underlined site is a HindIII site that was included in this primer and the bold-faced type is a NotI site). SEQ ID NO:65 is a reverse primer that anneals to 11340-11322 of HDMD in the 3' UTR. These primers cause the deletion of 707 bp of the 3' UTR from the XbaI cloning site located at 12057 to the end of this primer (SEQ ID NO:65), leaving 113 bp of native 3' UTR, and introducing NotI and HindIII cloning sites. The PCR product obtained using the primers corresponding to SEQ ID NOS:64 and 65 on the pΔR4-R23 clone was named HdysΔ3'UTR and was saved for use as a template to generate a further deletion of exons 71-78 (see part C below).

B. Deletion of 5' UTR and Alteration of Kozak Sequence

A portion of the 5' UTR was deleted (and the Kozak sequence was altered in the same step). The 'step 2' clone from cloning of ΔR4-R23 was utilized (this was the the product of ligating the step 1 PCR product into the 5016 bp NcoI and HindIII fragment from the HDMD full-length clone, and this clone contained pBSX backbone plus the 5' UTR, N terminus, Hinge 1, Repeats 1, 2, 3, Hinge 2, and part of repeat 24. There is an MunI site located in the first repeat at nucleotide 1409 of the HDMD cDNA. In addition, there is a NotI site that is polylinker derived at the 5' end of the clone. These two sites were employed, MunI + NotI, to clone a new fragment

containing a truncated 5' UTR and an altered Kozak sequence as follows. PCR was performed, using Pfu polymerase using the following primers. The first primer was 5' TAG CGG CCG CGG TTT TTT TTA TCG CTG CCT TGA TAT ACA CTT TCC ACC ATG CTT TGG TGG GAA GAA GTA G 3' (SEQ ID NO:59). We created a NotI site (underlined) in this primer so the product could be cloned back into the NotI site from the polylinker. The sequence immediately 3' to this NotI site corresponds to the dystrophin 5' UTR sequence (the original Kozak sequence was changed with this primer, from TCAAAATGC, changed to CCACCATGC. The second primer was 5' TTT TCC TGT TCC AAT CAG C 3' (SEQ ID NO:60) which anneals to sequence 1441-1423 of HDMD full length clone. The final product of this reaction was 1270 bp and was digested with NotI+MunI to produce a 1233 bp fragment that was then cloned into the NotI (polylinker) + MunI site in Repeat 1 of the "Step 2" clones (described above for ΔR4-23). This new clone was named pHDMD5' Kozak.

C. **Deletion of exons 71-78 (C-terminal)**

Using three PCR reactions, a 71-78 deletion was created. We used the HindIII fragment containing the 3'UTR that was generated by digestion of the HDMD full-length dystrophin cDNA with HindIII as the vector to clone the 71-78 fragment into the HindIII site. The primer employed for the first reaction were 5' GGC TTC CTA CAT TGT GTC AGT TTC CAT GTT GTC CCC 3' (SEQ ID NO:66), and 5' TCT CTC CAA GAT CAC CTC 3' (SEQ ID NO:67) anneals to 9117-9134 of HDMD. PCR was performed employing these primers and a 1334 bp product was produced. The primers for the second reaction were SEQ ID NO:65, and 5' **GGG GAC AAC ATG GAA ACT** GAC ACA ATG TAG GAA GCC 3' (SEQ ID NO:68), where the bold-face sequence anneals to exon 70 at 10415-10431 in the forward direction, and the underlined sequence anneals to 11216-11233 in the forward direction. PCR was performed and a 150 bp fragment was generated. The product of reactions 1 and 2 were used as template and the outside primers SEQ ID NO:65 and SEQ ID NO:67 were used to prime the reaction which generated the complete 71-78 C terminus (1484 bp). This product was digested with HindIII to produce a 1319 bp fragment and was

cloned into the HindIII site of pTZ19R (*See* Figure 35). This new clone was named pTZ-HDMD-H3Δ71-78Δ3.

D. Cloning of the SV40 pA Sequence into the NOt I site

The next step was the cloning of the SV 40 pA sequence: 5'GATCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGA ATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATT TGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATT TTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTCGGATC3' (SEQ ID NO:71) into the NotI site of the 3' HindIII fragment that now contains the 3' UTR and 71-78. A PCR reaction was performed on the template pHSA with a reverse primer 5' AGC GGC CGC AAA AAA CCT CCC ACA CCT CC 3' (SEQ ID NO:69, containing a regenerating NotI site - underlined) and 5' TAC **GGC CGA** TCC AGA CAT GAT AAG ATA C 3' (SEQ ID NO:70, containing a destroying EagI site, in bold). All other sequence (besides the NotI and EagI sites) is SV40 pA. This PCR reaction generated a 195 bp product + cloning sites = 209 bp. We then cloned this fragment into the NotI site of pTZ-HDMD-H3Δ71-78Δ3 generated by PCR in the 3' UTR clone. The upstream (5' - most) NotI site in this clone was destroyed by EagI ligation. This new clone was named pTZ-HDMD-H33'A.

E. Cloning of CK6 promoter into NotI site

The CK6 promoter - 5' GGT-ACTACGGGTCTAGGCTGCCCATGTAAGGAGGCAAGGCCTGGGGACACCCGAG ATGCCTGGTTATAATTAACCCCAACACCTGCTGCCCCCCCCCCCCCCAACACCT GCTGCCTGAGCCTGAGCGGTTACCCCACCCCGGTGCCTGGGTCTTAGGCTCTG TACACCATGGAGGAGAAGCTCGCTCTAAAAATAACCCTGTCCCTGGTGGGCC CAATCAAGGCTGTGGGGGACTGAGGGCAGGCTGTAACAGGCTTGGGGGCCA GGGCTTATACGTGCCTGGGACTCCCAAAGTATTACTGTTCCATGTTCCCGGCG AAGGGCCAGCTGTCCCCCGCCAGCTAGACTCAGCACTTAGTTTAGGAACCAG TGAGCAAGTCAGCCCTTGGGGCAGCCCATACAAGGCCATGGGGCTGGGCAAG

CTGCACGCCTGGGTCCGGGGTGGGCACGGTGCCCGGGCAACGAGCTGAAAGC
TCATCTGCTCTCAGGGCCCCTCCCTGGGGACAGCCCCTCCTGGCTAGTCACAC
CCTGTAGGCTCCTCTATATAACCCAGGGGCACAGGGGCTGCCCCCGGGTCAC
GGGGATCCTCTAGACC-3' (SEQ ID NO:61) was amplified using two tailed primers: 5' AGC GGC CGC GGT ACT ACG GGT CTA GG 3' Forward (SEQ ID NO:62), and 5' ATC GGC CGT CTA GAG GAT CCC CGT GAC C 3' Reverse (SEQ ID NO:63). The underlined sequence is a NotI site added to the end of this primer. The remaining sequence is CK6 sequence. The bold-faced type is an EagI site added to the end of this primer. The remaining sequence is from CK6. The CK6 promoter was amplified this way so we could add the NotI and EagI sites (so the entire cassette could be excised when put back together with NotI). This PCR product was therefore digested with NotI and EagI and ligated into the NotI site of pHDMD5'Kozak. This new clone was named pCK6HDMD5'Kozak. NotI and EagI produce compatible cohesive sites, but when EagI ligates to NotI, it destroys the site. So we placed the EagI site at the 3' end, so that when the final construct was cut with NotI, the entire expression cassette could be excised intact. The same strategy was employed at the 3' end when placing the SV40 poly A sequence into the 3' Not I site.

F. Re-ligating the 5' and 3' ends.

This step was performed as described above in the micro-dystrophin transgene constructs. We reconstituted the same cloning sites but with modifications in the fragments, so the modified 3' end, isolated as a HindIII fragment from clone pTZ-HDMD-H33'A (example 3 part D), was able to be cloned into the HindIII site of pCK6HDMD5'Kozak (example 3, part E). This final clone, named pCK6R4-R23KozakΔ3', contains a truncated dystrophin expression cassette that can be excised in its entirety by digestion with NotI. This excised expression cassette can then be used for a variety of purposes. One such purpose is to clone the cassette into a plasmid containing the inverted terminal repeats from adeno-associated virus. By cloning the dystrophin expression cassette HDMD-H33'A into a cloning site between the two ITRs of AAV, a recombinant AAV vector could be produced.

Example 4

Construction of Reduced Repeat Dystrophin Constructs

This example describes the construction of ΔH2-R19 (an 8 spectrin-like-repeat sequence), pΔR9R16 (a 16 spectrin-like-repeat sequence), pΔR1R24 (a zero spectrin-like-repeat sequence), pΔH2-H3 (a 8 spectrin-like repeat sequence), and ΔH2-R19,R20 (a 7 spectrin-like repeat sequence). One starting plasmid was pHBMD, a human dystrophin cDNA (full-length HDMD, SEQ ID NO:47) with a further deletion of the sequences encoded by exons 17-48. The cDNA was cloned into the commercially available plasmid vector pTZ19r (MBI Fermentas; Genbank accession number Y14835, See Fig.35), into which an EcoRI-SalI adapter (prepared by self-annealing of the oligonucleotide 5'-AATTCGTCGACG-3', SEQ ID NO:83) had been ligated into the the EcoRI site. Base number 1 of the cDNA is immediately 3' of the adapter sequence, and the cDNA ends at the XbaI site at base 12,100 of SEQ ID NO:1. This XbaI site had been ligated into the XbaI site of the plasmid ptZ19r. Another starting plasmid is pBSX (SEQ ID NO:46), a modified version of pBluescript KSII+ (Stratagene) which is used to make pBSXA (pBSX into which the SV40 polyadenylation signal (pA) was added). This pA sequence was excised as a 206 bp fragment from pCMVß (Clonetech), blunt-ended with DNA polymerase I, and ligated into the blunt-ended KpnI site of pBSX. Another starting plasmid is pCK3, which is pBSX with the 3.3 kb mouse muscle creatine kinase enhancer plus promoter attached to the minx intron (*See*, Hauser *et al.*, *Mol Ther.*, 2:16-25, 2000). Another staring plasmid is pHDSK, which is pHBMD digested with KpnI, to remove the dystrophin sequences 3' of the internal KpnI site (base 7,616 of the human dystrophin cDNA sequence, SEQ ID NO:1). A further starting vector is p44.1, which is pBluescript KS- (Stratagene) carrying a human dystrophin cDNA fragment spanning the EcoRI site at base 7,002 to the EcoRI site at base 7,875 of the full-length human dystrophin cDNA sequence, cloned into the EcoRI site of the vector. Another plasmid employed was p30-2, pBluescribe (Stratagene) containing a fragment from the full-length human dystrophin cDNA spanning bases 1,455 to the EcoRI site at base 2,647, cloned into the EcoRI site of the vector. An additional vector employed was p30-1, pBluescribe

(Stratagene) containing an EcoRI fragment from the full-length human dystrophin cDNA spanning bases 2,647 to 4,558, cloned into the EcoRI site of the vector. An further plasmid employed is p47-4, pBluescript KS- (Stratagene) carrying the human dystrophin cDNA EcoRI fragment spanning bases 4,452 to 7,002 of the full-length cDNA sequence, cloned into the EcoRI site of the vector. Another plasmid is p9-7, pBluescribe (Stratagene) containing bases 1-1,538 of the full-length human dystrophin cDNA. Base 1 is attached to a linker of the sequence 5' GAATTC-3' and cloned into the EcoRI site of the vector. Base 1,538 is blunt-end cloned into the PstI site of the vector, which had been destroyed by fill-in with T4 DNA polymerase. Another vector employed is p63-1, pBluescript KS- (Stratagene) carrying the human dystrophin cDNA EcoRI fragment spanning bases 7,875 to the 3' end of the full-length cDNA, cloned into the EcoRI site of the vector (the 3' end of the cDNA is ligated to a linker of the sequence 5'-GAATTC-3').

Initially, the MCK promoter plus enhancer and the minx intron were excised from pCK3 by digestion with EagI, yielding a 3.5 kb fragment that was ligated into EagI-digested pBSXA to make pBSXACK3. Truncated dystrophin cDNAs, derived from pHBMD, containing various deletions of dystrophin domains were prepared as described below. The cDNA inserts were excised from the plasmid backbone with SalI, and ligated into pBSXACK3 at the SalI site, which is located between the minx intron and the pA sequence, such that the 3' end of the cDNA was adjacent to the pA sequence. The isolation of the truncated cDNAs is described below. pBSXACK3-truncated dystrophin plasmids were digested with BssHII to release the expression vectors, which were gel purified and used to generate transgenic mice.

Isolation of ΔH2R19

A PCR product was generated by amplification of plasmid p30-2 with primers 5'-TGTGCTGCAAGGCGATTAAGTTGG-3' (SEQ ID NO:72) and 5'-GAGCTAGGTCAGGCTGCTGTGAAATCTGTGC-3' (SEQ ID NO:75). Primer SEQ ID NO:75 overlaps the end of repeat 3 and the beginning of hinge 3. Primer SEQ ID NO:72 corresponds to a sequence in the plasmid vector adjacent to the

cloning site. A second PCR product was generated by amplification of plasmid p44-1 using primers 5'-CCAGGCTTTACACTTTATGCTTCC-3' (SEQ ID NO:73) and 5'-GCACAGATTTCACAGCAGCCTGACCTAGCTC-3' (SEQ ID NO:74). Primer SEQ ID NO:74 is the reverse complement of primer SEQ ID NO:75. Primer SEQ ID NO:73 corresponds to a sequence in the plasmid vector adjacent to the cloning site. The PCR products were then purified by agarose gel electorphoreses, and quantified. A recombinant PCR product was then prepared by mixing together 10 ng of each of the first two PCR products, then re-PCR amplifying using only primers SEQ ID NO:72 and SEQ ID NO:73. This recombinant PCR product was then digested with NheI and KpnI, and ligated into NheI and KpnI digested pHΔSK to generate plasmid pHBMDΔH2 (NheI cuts at cDNA base 1,519, and KpnI cuts at base 7,616 of the full-length human dystrophin cDNA sequence). pHBMDΔH2 was then digested with KpnI and Xba1, and ligated to the KpnI-XbaI fragment from pHBMD (this latter fragment contains the full-length human dystrophin cDNA bases 7,616 to 12,100) to obtain plasmid pΔH2R19.

Isolation of pΔR9R16

Plasmid p44-1 was digested with EcoRI and Asp718 to excise a 610 bp cDNA insert, that was ligated into pBSX digested with EcoRI and Asp718, yielding pBSX44AE. pBSX44AE was digested with EcoRI and XbaI, and ligated to the NheI-EcoRI cDNA-containing fragment from p30-2, yielding pBSX44AE/30-2NE. Plasmid pBSX44AE/30-2NE was linearized by digestion with EcoRI, into which was ligated the EcoRI-digested recombinant PCR product ΔR9-R16. This latter recombinant PCR product was generated as follows. Plasmid p30-1 was amplified with primers SEQ ID NO:72 and 5'-CCATTTCTCAACAGATCTTCCAAAGTCTTG-3' (SEQ ID NO:77), and plasmid p47-4 was amplified by PCR with primers SEQ ID NO:73 and 5'-CAAGACTTTGGAAGATCTGTTGAGAAATGG-3 (SEQ ID NO:76). A recombinant PCR product (ΔR9-R16) was then prepared by mixing together 10 ng of each of the first two PCR products, then re-PCR amplifying using only primers SEQ ID NO:72 and SEQ ID NO:73. This recombinant PCR product was then

digested with EcoRI, and ligated into EcoRI digested pBSX44AE/30-2NE to generate plasmid pR9R16int. Plasmid pR9R16int was digested with NcoI and Asp718, and the 3 kb cDNA fragment was isolated and ligated into NcoI and Asp718 digested pHΔSK to generate pΔR9R16.

Isolation of pΔR1R24

Plasmid p9-7 was PCR amplified with PCR primers 5'-AGTGTGGTTTGCCAGCAGTC (SEQ ID NO:80) and 5'-CAAAGTCCCTGTGGGCGTCTTCAGGAGCTTCC-3' (SEQ ID NO:79). Plasmid p63-1 was PCR amplified with primers 5' GGAAGCTCCTGAAGACGCCCACAGGGACTTTG-3' (SEQ ID NO:78) and 5'-TGGTTGATATAGTAGGGCAC-3' (SEQ ID NO:81). A recombinant PCR product (ΔR1-R24) was then prepared by mixing together 10 ng of each of the first two PCR products, then re-PCR amplifying using only primers SEQ ID NO:80 and SEQ ID NO:81. This recombinant PCR product was then digested with SexAI and PpuMI, and ligated into SexAI and PpuMI digested pHBMD to generate plasmid pΔR1R24.

Isolation of pΔH2-H3

This clone was prepared exactly as pΔH2-R19, except that primer 5'-CAGATTTCACAGGCTGCTCTGGCAGATTTC-3' (SEQ ID NO:82) was used in place of primer SEQ ID NO:74, and primer 5'-GAAATCTGCCAGAGCAGCCTGTGAAATCTG-3' (SEQ ID NO:84) was used in place of primer SEQ ID NO:75.

Isolation of ΔH2-R19,R20

This clone was generated from clone pΔH2R19 as follows. Plasmid p44-1 was amplified with primers SEQ ID NO:72 and 5'-TGAATCCTTTAACATAGGTACCTCCAACAT-3' (SEQ ID NO:85). Plasmid 63-1 was amplified with primers 5'-ATGTTGGAGGTACCTATGTTAAAGGATTCA-3' (SEQ ID NO:86) and SEQ ID NO:81. The PCR products were then purified by

agarose gel electorphoreses, and quantified. A recombinant PCR product was then prepared by mixing together 10 ng of each of the first two PCR products, then re-PCR amplifying using only primers SEQ ID NO:72 and SEQ ID NO:81. This product was digested with Asp718 and BstXI, and ligated into Asp718 and BstXI digested pHBMD generating clone pBMDΔR20. The Asp718-XbaI cDNA-containing fragment from pBMDΔR20 was isolated and ligated into Asp718 and XbaI digested pΔH2R19 to generate pΔH2-R19,R20.

EXAMPLE 5

Testing Truncated Dystrophin in *mdx* Mice

This example describes the generation of transgenic mdx mice expressing truncated dystrophin cDNA (see above), and testing these mice in various ways to determine various measurable muscle values. A variety of dystrophin expression cassettes (Fig. 27) were used to generate transgenic mice to test their functional capacity in alleviating muscular dystrophy on the dystrophin null *mdx* background. Figure 27 depicts the truncated dystrophin cDNA sequences tested, all of which were linked to an regulatory regions, a minx intron, and the SV40 polyadenylation sequence (the 4-repeat constructs employed the HSA actin promoter, *See* Crawford *et al., J. Cell. Biol.*, 150:1399, 2000; and the remaining sequences employed an MCK enhancer and promoter, see Niwa *et al., Genes Dev.* 4:1552, 1990). Each of these constructs was released by digestion from plasmid hosts, were gel purified, and used to generate transgenic mice.

Excised expression cassettes injected into wild type C57Bl/10 x SJL/J F2 hybrid embryos, and F^0 mice were screened by PCR analysis of DNA isolated from tail snips. Positive F^0 mice were backcrossed onto the C57Bl/10*mdx* background, and individual mouse lines were tested for dystrophin expression by immunofluorescent analysis with dystrophin antibodies for of expression in skeletal muscle fibers. Lines that displayed uniform expression of dystrophin in muscle fibers were selected for further analysis. These lines were further backcrossed onto the *mdx* mouse

background before analysis of dystrophin expression, muscle function and morphology.

A. Truncated dystrophin cDNAs are expressed at various levels in muscles of transgenic mdx mice.

Muscle extracts were analyzed by western (immuno) blot analysis to determine the amount of dystrophin made in different muscles of the transgenic *mdx* mice. For these studies, total protein was extracted from the quadriceps and diaphragm muscles of control and transgenic mice, and protein concentrations were determined using the Coomassie Plus Protein Assay Reagent (Pierce). One hundred micrograms of each sample was electrophoresed on a 6% polyacrylamide/SDS gel (29.7:0.3/acryl:bis), transferred for 2 hours at 75 volts onto Biotrace Nitrocellulose (Gelman Science) in 1X Tris-Glycine, 20% methanol, 0.05% SDS, using a wet-transfer apparatus (Hoefer). Membranes were blocked in 10% non-fat dry milk, 1% normal goat serum, and 0.1% Tween-20, and hybridized with DYS1 (Novacastra) at a 1/1000 dilution for 2 hours at room temperature, washed, and then probed with horse radish peroxidase conjugated anti-mouse antibodies at a 1/2,000 dilution (Cappel). Blots were developed using the ECL chemiluminescence system (Amersham). All incubations contained 1% normal goat serum and 0.1% Tween-20. The results of the western blot indicated that R9-R16 was poorly expressed in this line of mice, especially in the diaphragm, and that H2-H3 was very poorly expressed in the diaphragm.

B. Truncated dystrophin cDNAs confer various degrees of protection on muscles of transgenic mdx mice.

Various muscle groups from the different lines of transgenic mice expressing truncated dystrophins were examined for morphological abnormalities, and for expression of dystrophin by indirect immunofluorescence (IF) in individual fibers. IF analysis was performed as follows. Skeletal muscle was removed from control and transgenic animals, cut into strips, embedded in Tissue-tek OCT mounting media (Miles, Inc.), and frozen quickly in liquid nitrogen-cooled isopentane. Seven micrometer sections were blocked with 1% gelatin in KPBS for 15 minutes, washed in

KPBS + 0.2% gelatin (KPBSG), and incubated for 2 hours in KPBSG + 1% normal goat serum with affinity-purified dystrophin antibody 18-4 (Cox *et al.*, *Nature*, 364:725-729, 1993) at a dilution of 1/1000. After washing, the slides were incubated for 1 hour with either biotin-labeled goat anti-rabbit polyclonal antibodies (Pierce), washed again, and incubated with FITC (fluorescein isothiocynate)-conjugated streptavidin. After a final wash, Vectashield (Vector Laboratories, Inc.) with DAPI was applied and sections were photographed through a dual bandpass filter under 40X magnification using a Nikon E1000 microscope.

Morphological analysis of the muscles was performed as follows. Muscle groups from among the following types were chosen for analysis: Quadriceps (Quad), soleus, extensor digitorum longus (EDL), tibialis anterior (TA), and diaphragm muscles. Selected muscles were removed from mice, frozen in liquid nitrogen cooled O.C.T. embedding medium (Tissue-Tek), and cut into 7 μm sections. After fixing in 3.7% formaldehyde, sections were stained in hematoxylin and eosin-phloxine. Stained sections were imaged with a Nikon E1000 microscope and photographed.

The results of this analysis show that micro-dystrophin expression (ΔR4R23 transgene) in the diaphragm prevents the onset of muscular dystrophy in *mdx* mice. In particular, micro-dystrophin transgenic and wild-type C57Bl/10 diaphragm sections stained with hematoxylin and eosin (H&E) show morphologically healthy muscle without areas of fibrosis, necrosis, mononuclear cell infiltration, or centrally located nuclei. Conversely, the *mdx* diaphragm displays a high level of dystrophic morphology by H&E. Also, immuno-fluorescence, using anti-dystrophin polyclonal primary antisera, demonstrates that micro-dystrophin transgenes are expressed at the sarcolemmal membrane in a similar fashion to that of wild-type dystrophin, while *mdx* mice do not express dystrophin.

H & E staining also shows that truncated dystrophins with 8 or 16 spectrin-like repeats have varying abilities to prevent dystrophy in the diaphragm of transgenic *mdx* mice. The H2R19 maintains normal muscle morphology that is not different from wild-type C57Bl/10 muscle. The ΔH2R19 muscle displays a very low percentage of

centrally nucleated fibers, while the ΔH2-R19,R20 and ΔR9-16 constructs display percentages intermediate between ΔH2-R19 and *mdx* (see Fig. 28). The *mdx* diaphragm had a large number of centrally nucleated fibers, many necrotic fibers, and large areas of mono-nuclear cell infiltration and fibrosis.

The results also show that quadriceps muscle fibers expressing micro-dystrophin transgene (ΔR4R23 transgene) display normal morphology and exclude Evans Blue Dye. Micro-dystrophin transgenic *mdx* or C57Bl/10 quadriceps sections stained with hematoxylin and eosin (H&E) display morphologically healthy muscle without areas of necrosis, fibrosis, mononuclear cell infiltration, or centrally-located nuclei, as opposed to sections of *mdx* muscle. The high abundance of central nuclei and mononuclear immune cell infiltration are evidence of muscle cell necrosis. Immunofluorescence results indicate that micro-dystrophins display a subsarcolemmal expression pattern like that of wild-type dystrophin, while *mdx* mice do not express dystrophin. Evans Blue Dye (EBD) uptake is an indication of a damaged myofiber. For analysis of EBD uptake, mice were tail vein injected with 150 µl of a solution containing 10 mg/ml Evans blue dye in PBS (150 mM NaCl, 50 mM Tris pH 7.4). After three hours, the animals were euthanized and mouse tissues were either fixed in 3.7% formaldehyde/0.5% glutaraldehyde to observe gross dye uptake, or frozen unfixed in O.C.T. embedding medium. To examine Evans blue uptake by individual fibers, 7 µm thick frozen sections were fixed in cold acetone and analyzed by fluorescence microscopy. The results of this testing indicate that fibers expressing micro-dystrophin or wild-type dystrophin exclude EBD, and that damaged *mdx* muscle cell membranes are permeable to Evans Blue dye.

A hallmark of dystrophy in *mdx* mice is the presence of large numbers of centrally-nucleated muscle fibers, reflecting cycles of fiber degeneration and regeneration. To estimate the degree of myofiber regeneration occurring in the transgenic mice, centrally-nucleated fibers were counted from quadriceps muscles in age-matched wild-type, *mdx*, and transgenic mdx mice (Fig. 28). To determine the

percentage of fibers containing central nuclei, the number of muscle fibers with centrally-located nuclei was divided by the total number of muscle fibers.

Expression of 8 or 4 repeat micro-dystrophin transgenes on the *mdx* background significantly reduces the percentage of fibers with centrally-located nuclei to wild-type or near wild-type levels (Fig. 28). Dystrophin molecules with zero repeats are unable to correct the *mdx* phenotype by this assay. The best constructs were observed to be the 8 repeat H2-R19 and the 4 repeat R2-R23 constructs. Greater percentages of centrally nucleated fibers were observed in mice expression the exon 17-48 deletion, the 4 repeat R2R21 construct, the 7 repeat H2R19,R20 construct, the 16 repeat R9R16 construct, and the zero repeat R1R24 construct (Fig. 28). The results from the R9R16 construct likely do not reflect the full functional capacity of the 16 repeat dystrophin since this line of mice expressed very low levels of the truncated dystrophin protein. All other muscles expressed levels of dystrophin that have been shown to be capable of preventing dystrophy if the expressed protein is functional (Phelps *et al., Hum Mol Genet*; 4:1251-1258, 1995).

The functional capacity of the truncated dystrophins was also assessed by measuring muscle contractile properties in the transgenic *mdx* mice. Contractile properties of muscles from transgenic mice were compared with those of C57Bl/10 wild type and *mdx* mice. The samples included 4-8 muscles each from the tibialis anterior (TA), extensor digitorum longus (EDL) or diaphragm. Mice were deeply anesthetized with avertin and each muscle was isolated and dissected free from the mouse. After removal of the limb muscles, the mice were euthanized with the removal of the diaphragm muscle. The muscles were immersed in a bath filled with oxygenated buffered mammalian Ringer's solution (137 mM NaCl, 24 mM $NaHCO_3$, 11 mM glucose, 5 mM KCl, 2 mM $CaCl_2$, 1 mM $MgSO_4$, 1 mM NaH_2PO_4, and 0.025 mM tubocurarine chloride, pH 7.4). For each muscle, one tendon was tied to a servomotor and the other tendon to a force transducer. Muscles were stretched from slack length to the optimal length for force development and then stimulated at a frequency that produced absolute isometric tetanic force (mN). Following the

measurements of the contractile properties, the muscles were removed from the bath, blotted and weighed to determine muscle mass. Specific force (kN/m2) was calculated by dividing absolute force by total fiber cross sectional area.

Figure 29 shows that the 8 repeat dystrophin encoded by H2-R19 supports normal force development in both the diaphragm (Fig. 29a) and EDL muscle (Fig. 29b). In contrast, previous studies showed that the exon 17-48 construct, which encodes a dystrophin with 8.25 spectrin-like repeats, supports only 90-95% of normal force development in the diaphragm (Phelps *et al., Hum Mol Genet,* 4:1251-1258, 1995). The 8 repeat dystrophin lacking a central hinge (H2-H3), and the 7 repeat dystrophin (H2-R19,R20) both fail to support significant force generation compared with dystrophic *mdx* muscles. The results from the R9-R16 construct likely do not reflect the full functional capacity of the 16 repeat dystrophin, since this line of mice expressed very low levels of the truncated dystrophin.

Figure 30 shows that the micro-dystrophin transgenic *mdx* mice develop less specific force than do C57Bl/10 mice in the TA, but near wild-type levels in the diaphragm. Micro-dys 1 and -2 refer to transgenes ΔR4-R23, and ΔR2-R21, respectively. Figure 30A shows that C57Bl/10 mice display significantly higher specific force than both transgenic lines and mdx mice in the tibialis anterior (TA) muscle. Data are presented as means ± standard error of the means (s.e.m.) with each bar representing 6 to 8 TA muscles. ANOVA statistical testing was performed. (* indicates significance from C57Bl/10, $p<0.01$; s indicates significance from C57Bl/10, $p<0.05$). Figure 30B shows that mice expressing Micro-dys 1 develop wild type levels of specific force in the diaphragm, while mice expressing Micro-dys 2 develop ~22% less specific force by the same assay when compared with C57Bl/10. Both lines of mice develop more specific force than mdx mice in the diaphragm. Data are presented as the percentage of wild type.

Dystrophic mice are susceptible to contraction-induced injury (Petrof, *et al., Proc. Natl. Acad. Sci.* USA. 90:3710-3714, 1993). In this part of the example tested whether the 4 repeat dystrophin clones would protect muscles of transgenic *mdx* mice

from contraction induced injuries. To test contraction-induced injury, an experimental protocol consisting of two muscle stretches was performed in live, anesthetized animals. The distal tendon of the TA was cut and secured to the lever arm of a servomotor that monitors position and force produced by the muscle. Stimulation voltage and optimal muscle length (L_0) for force production were determined. The muscle was maximally stimulated and then stretched 40% greater than L_0 (LC1) for 300 milliseconds. A second lengthening contraction was performed 10 seconds later (LC2). The maximum force that the muscle was able to produce after each stretch was measured and expressed as a percentage of the force produced before stretch. *Mdx* mice expressing micro-dystrophins were significantly protected from the dramatic force deficit produced after a lengthening contraction compared with mdx mice (Fig. 31). Micro-dys 1 and -2 refer to transgenes ΔR4-R23, and ΔR2-R21, respectively. Furthermore, there was no significant difference between either micro-dystrophin construct studied in this assay and C57Bl/10 mice following the second, most damaging lengthening contraction. Data are presented as means ± s.e.m. with each bar representing between 6 and 8 TA muscles from 9-11 week old mice.

C. Truncated 4 repeat dystrophin cDNAs restore the ability to run long distances to mdx mice.

We have observed that *mdx* mice are not able to run for long distances on a treadmill, as compared to wild-type mice (*see* below). Therefore, mice expressing four repeat dystrophins were compared with wild-type and *mdx* mice for ability to run for extended times on a treadmill. The exercising protocol utilized a six lane, enclosed treadmill with a shock grid to allow forced running at a controlled rate. C57Bl/10, C57Bl/6 x SJL F1, *mdx* or transgenic *mdx* mice were run at a 15 degree downward angle to induce damaging eccentric muscle contractions. Mice were given a 15 minute acclimation period prior to exercise, and then ran at 10 meters/minute with a subsequent 5 m/min increase in rate every 10 minutes until exhaustion. Exhaustion was determined to be the time at which a mouse spent more than 5 seconds sitting on

the shock grid without attempting a re-entry to the treadmill. As shown in Fig 32, both lines of four repeat transgenic mice ran significantly farther than *mdx* mice. Micro-dys 1 and -2 refer to transgenes ΔR4-R23, and ΔR2-R21, respectively. Micro-dystrophin transgenic mice are a genetic mixture of C57Bl/6 x SJL, and C57Bl/10 strains, and ran an intermediate distance between the two wild-type lines. Data are presented as means ± s.e.m. ANOVA statistical analyses were performed. (* indicates values significantly different from mdx line, $p<0.01$; s indicates values significantly different from mdx line, $p<0.05$).

D. Micro-dystrophin transgenic *mdx* mice do not display hypertrophy

As a way to measure the functional capacity of the four-repeat dystrophins, we weighed both whole mice and dissected tibialis anterior muscles from age matched transgenic and control mice. The results shown in Fig. 33 show that the micro-dystrophin transgenic mdx mice do not display the muscle hypertrophy normally observed in *mdx* mice. Figure 33A shows that three month old micro-dystrophin transgenic *mdx* mice weighed significantly less than age-matched *mdx* control mice. Figure 11B shows that tibialis anterior (TA) muscle masses in *mdx* mice were significantly higher than control muscle masses in C57Bl/10 and in both lines of *mdx* mice expressing different micro-dystrophin transgenes. Data are presented as means ± s.e.m. with each bar representing between 3 and 4 mice. ANOVA statistical analyses were performed (* indicates difference from mdx line, $p<0.01$; Y indicates difference from C57Bl/10 line, $p<0.01$; s indicates difference from C57Bl/10 line, $p<0.05$). Micro-dys 1 and -2 refer to transgenes ΔR4-R23, and ΔR2-R21, respectively.

EXAMPLE 6

Microdystrophin-containing Adeno-associated Viral Vectors

This example describes a construct that could be made in order to allow adeno-associated virus to express a mini-dystrophin peptide in a target muscle cells. Fig. 34 shows a schematic illustration of a plasmid vector containing the adeno-associated

virus inverted terminal repeats (AAV-ITRs), the muscle promoter plus enhancer fragment known as CK6 (SEQ ID NO:61, the ΔR2-R21 four repeat dystrophin cDNA (SEQ ID NO:40) with a further deletion of sequences encoded on exons 71-78, plus a 195 base pair SV40 polyadenylation signal that would have a total insert size of approximately 4.7 kb. The cloning capacity of adeno-associated viral vectors is approximately 4.9 kb. As such, the construct could be efficiently packaged into AAV viral particles (*e.g.* this plasmid construct could be used to transfect cells such that AAV expressing mini-dystrophin peptide is expressed). These AAV then, for example, may be administered to a subject with DMD or BMD (*i.e.* gene therapy to correct a muscle deficiency in a subject).

All publications and patents mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in material science, chemistry, and molecular biology or related fields are intended to be within the scope of the following claims.

CLAIMS

We claim:

1. A composition comprising nucleic acid encoding a mini-dystrophin peptide, wherein said mini-dystrophin peptide comprises a spectrin-like repeat domain, and wherein said spectrin-like repeat domain consists of n spectrin-like repeats, wherein n is an even number less than 24.

2. The composition of Claim 1, wherein said mini-dystrophin-peptide is capable of altering a measurable muscle value in a DMD animal model by at least 20% of the wild type value.

3. The composition of Claim 1, wherein said mini-dystrophin peptide is capable of altering a measurable muscle value in a DMD animal model to a level similar to the wild-type value.

4. The composition of Claim 1, wherein n is a multiple of 4.

5. The composition of Claim 1, wherein n is 4.

6. The composition of Claim 1, wherein said nucleic acid comprises an expression vector.

7. The composition of Claim 1, wherein said nucleic acid comprises spectrin-like repeat encoding sequences.

8. The composition of Claim 7, wherein said spectrin-like repeat encoding sequences are precise spectrin-like repeat encoding sequences.

9. The composition of Claim 1, wherein said nucleic acid comprises an actin-binding domain encoding sequence.

10. The composition of Claim 9, wherein said actin binding domain comprises at least a portion of SEQ ID NO:6.

11. The composition of Claim 1, wherein said nucleic acid comprises a β-dystroglycan binding domain.

12. The composition of Claim 11, wherein said β-dystroglycan binding domain comprises at least a portion of a dystrophin hinge 4 encoding sequence, and at least a portion of a dystrophin cysteine-rich domain encoding sequence.

13. The composition of Claim 7, wherein said spectrin-like repeat encoding sequences are selected from the group consisting of SEQ ID NOS:8-10, 12-27, and 29-33.

14. The composition of Claim 1, wherein said nucleic acid contains less than 75% of a wild type dystrophin 3' untranslated region.

15. The composition of Claim 1, wherein said mini-dystrophin peptide further comprises a substantially deleted dystrophin C-terminal domain.

16. The composition of Claim 1, wherein said nucleic acid sequence contains less than 50% of a dystrophin 3' untranslated region.

17. A method of expressing a mini-dystrophin peptide in a target cell, comprising;

a) providing;

i) a vector comprising nucleic acid encoding a mini-dystrophin peptide, wherein said mini-dystrophin peptide comprises a spectrin-like repeat domain, and wherein said spectrin-like repeat domain consists of n spectrin-like repeats, wherein n is an even number less than 24, and

ii) a target cell, and

b) contacting said vector with said target cell under conditions such that said mini-dystrophin peptide is expressed in said target cell.

18. The method of Claim 17, wherein said mini-dystrophin peptide further comprises a substantially deleted dystrophin C-terminal domain.

19. The composition of Claim 17, wherein said nucleic acid comprises spectrin-like repeat encoding sequences.

20. The method of Claim 19, wherein said spectrin-like repeat encoding sequences are precise spectrin-like repeat encoding sequences.

21. The composition of Claim 17, wherein said mini-dystrophin-peptide is capable of altering a measurable muscle value in a DMD animal model by at least 20% of the wild type value.

22. The composition of Claim 17, wherein said mini-dystrophin peptide is capable of altering a measurable muscle value in a DMD animal model to a level similar to the wild-type value.

23. A composition comprising nucleic acid, wherein said nucleic acid encodes a mini-dystrophin peptide, and wherein said mini-dystrophin peptide comprises a substantially deleted dystrophin C-terminal domain.

24. The composition of Claim 23, wherein said substantially deleted dystrophin C-terminal domain is less than 40% of a wild type dystrophin C-terminal domain.

25. The composition of Claim 23, wherein said mini-dystrophin-peptide is capable of altering a measurable muscle value in a DMD animal model by at least 20% of the wild type value.

26. The composition of Claim 23, wherein said mini-dystrophin peptide is capable of altering a measurable muscle value in a DMD animal model to a level similar to the wild-type value.

27. The composition of Claim 23, wherein said nucleic acid comprises an actin-binding domain encoding sequence.

28. The composition of Claim 23, wherein said nucleic acid comprises a β-dystroglycan binding domain.

29. The composition of Claim 23, wherein said nucleic acid comprises at least 2 spectrin-like repeat encoding sequences.

30. The composition of Claim 23, wherein said nucleic acid comprises viral nucleic acid.

PATENT
Attorney Docket No.: UM-04723

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of: Jeffrey S. Chamberlain *et al.* Group No.:
Serial No.: Examiner:
Filed:
Entitled: **Truncated Dystrophin Genes**

POWER OF ATTORNEY BY ASSIGNEE

Assistant Commissioner for Patents
Washington, D.C. 20231

Regents of the University of Michigan, as Assignee of record of the entire interest of the above-identified patent application, hereby appoints the members of the firm of MEDLEN & CARROLL, LLP, a firm composed of:

Virginia S. Medlen	(Reg. No. 32,050)	Maha A. Hamdan	(Reg. No. 43,655)
Peter G. Carroll	(Reg. No. 32,837)	J. Mitchell Jones	(Reg. No. 44,174)
Kamrin T. MacKnight	(Reg. No. 38,230)	David J. Wilson	(Reg. No. 45,225)
David A. Casimir	(Reg. No. 42,395)	Jason R. Bond	(Reg. No. 45,439)
)		Jaen Andrews	(Reg. No. 35,051)

as its attorneys with full power of substitution to prosecute this application and transact all business in the Patent and Trademark Office in connection therewith.

Please direct all future correspondence and telephone calls regarding this application to:

Jason R. Bond
MEDLEN & CARROLL, LLP
220 Montgomery Street, Suite 2200 Telephone: 415/705-8410
San Francisco, California 94104 Facsimile: 415/397-8338

I hereby certify that the Assignment document filed with the application or filed subsequent to the filing date of the application, has been reviewed and I hereby certify that, to the best of my knowledge and belief, title is with Regents of the University of Michigan.

Dated: _____ By: _____

Name: _____

Title: _____

Regents of the University of Michigan
3003 So. State Street
Ann Arbor, MI 48109

FIGURE 1 (Human Dystrophin cDNA, Acc. No. M18533, SEQ ID NO:1)

```
   1 gggattccct cactttcccc ctacaggact cagatctggg aggcaattac cttcggagaa
  61 aaacgaatag gaaaaactga agtgttactt tttttaaagc tgctgaagtt tgttggtttc
 121 tcattgtttt taagcctact ggagcaataa agtttgaaga acttttacca ggtttttttt
 181 atcgctgcct tgatatacac ttttcaaaat gctttggtgg gaagaagtag aggactgtta
 241 tgaaagagaa gatgttcaaa agaaaacatt cacaaaatgg gtaaatgcac aattttctaa
 301 gtttgggaag cagcatattg agaacctctt cagtgacctа caggatggga ggcgcctcct
 361 agacctcctc gaaggcctga cagggcaaaa actgccaaaa gaaaaaggat ccacaagagt
 421 tcatgccctg aacaatgtca acaaggcact gcgggttttg cagaacaata atgttgattt
 481 agtgaatatt ggaagtactg acatcgtaga tggaaatcat aaactgactc ttggtttgat
 541 ttggaatata atcctccact ggcaggtcaa aaatgtaatg aaaaatatca tggctggatt
 601 gcaacaaacc aacagtgaaa agattctcct gagctgggtc cgacaatcaa ctcgtaatta
 661 tccacaggtt aatgtaatca acttcaccac cagctggtct gatggcctgg ctttgaatgc
 721 tctcatccat agtcataggc cagacctatt tgactggaat agtgtggttt gccagcagtc
 781 agccacacaa cgactggaac atgcattcaa catcgccaga tatcaattag gcatagagaa
 841 actactcgat cctgaagatg ttgataccac ctatccagat aagaagtcca tcttaatgta
 901 catcacatca ctcttccaag ttttgcctca acaagtgagc attgaagcca tccaggaagt
 961 ggaaatgttg ccaaggccac ctaaagtgac taaagaagaa cattttcagt tacatcatca
1021 aatgcactat tctcaacaga tcacggtcag tctagcacag ggatatgaga gaacttcttc
1081 ccctaagcct cgattcaaga gctatgccta cacacaggct gcttatgtca ccacctctga
1141 ccctacacgg agcccatttc cttcacagca tttggaagct cctgaagaca agtcatttgg
1201 cagttcattg atggagagtg aagtaaacct ggaccgttat caaacagctt tagaagaagt
1261 attatcgtgg cttctttctg ctgaggacac attgcaagca caaggagaga tttctaatga
1321 tgtggaagtg gtgaaagacc agtttcatac tcatgagggg tacatgatgg atttgacagc
1381 ccatcagggc cgggttggta atattctaca attgggaagt aagctgattg gaacaggaaa
1441 attatcagaa gatgaagaaa ctgaagtaca agagcagatg aatctcctaa attcaagatg
1501 ggaatgcctc agggtagcta gcatggaaaa acaaagcaat ttacatagag ttttaatgga
1561 tctccagaat cagaaactga aagagttgaa tgactggcta acaaaaacag aagaaagaac
1621 aaggaaaatg gaggaagagc ctcttggacc tgatcttgaa gacctaaaac gccaagtaca
1681 acaacataag gtgcttcaag aagatctaga acaagaacaa gtcagggtca attctctcac
1741 tcacatggtg gtggtagttg atgaatctag tggagatcac gcaactgctg ctttggaaga
1801 acaacttaag gtattgggag atcgatgggc aaacatctgt agatggacag aagaccgctg
1861 ggttctttta caagacatcc ttctcaaatg gcaacgtctt actgaagaac agtgcctttt
1921 tagtgcatgg ctttcagaaa aagaagatgc agtgaacaag attcacacaa ctggctttaa
1981 agatcaaaat gaaatgttat caagtcttca aaaactggcc gttttaaaag cggatctaga
2041 aaagaaaaag caatccatgg gcaaactgta ttcactcaaa caagatcttc tttcaacact
2101 gaagaataag tcagtgaccc agaagacgga agcatggctg gataactttg cccggtgttg
2161 ggataattta gtccaaaaac ttgaaaagag tacagcacag atttcacagg ctgtcaccac
2221 cactcagcca tcactaacac agacaactgt aatggaaaca gtaactacgg tgaccacaag
2281 ggaacagatc ctggtaaagc atgctcaaga ggaacttcca ccaccacctc cccaaaagaa
2341 gaggcagatt actgtggatt ctgaaattag gaaaaggttg gatgttgata taactgaact
2401 tcacagctgg attactcgct cagaagctgt gttgcagagt cctgaatttg caatctttcg
2461 gaaggaaggc aacttctcag acttaaaaga aaaagtcaat gccatagagc gagaaaaagc
2521 tgagaagttc agaaaactgc aagatgccag cagatcagct caggccctgg tggaacagat
2581 ggtgaatgag ggtgttaatg cagatagcat caaacaagcc tcagaacaac tgaacagccg
2641 gtggatcgaa ttctgccagt tgctaagtga gagacttaac tggctggagt atcagaacaa
2701 catcatcgct ttctataatc agctacaaca attggagcag atgacaacta ctgctgaaaa
2761 ctggttgaaa atccaaccca ccacccccatc agagccaaca gcaattaaaa gtcagttaaa
2821 aatttgtaag gatgaagtca accggctatc aggtcttcaa cctcaaattg aacgattaaa
2881 aattcaaagc atagccctga aagagaaagg acaaggaccc atgttcctgg atgcagactt
2941 tgtggccttt acaaatcatt ttaagcaagt cttttctgat gtgcaggcca gagagaaaga
3001 gctacagaca atttttgaca ctttgccacc aatgcgctat caggagacca tgagtgccat
3061 caggacatgg gtccagcagt cagaaaccaa actctccata cctcaactta gtgtcaccga
3121 ctatgaaatc atggagcaga gactcgggga attgcaggct ttacaaagtt ctctgcaaga
3181 gcaacaaagt ggcctatact atctcagcac cactgtgaaa gagatgtcga agaaagcgcc
3241 ctctgaaatt agccggaaat atcaatcaga atttgaagaa attgagggac gctggaagaa
3301 gctctcctcc cagctggttg agcattgtca aaagctagag gagcaaatga ataaactccg
3361 aaaaattcag aatcacatac aaaccctgaa gaaatggatg gctgaagttg atgtttttct
3421 gaaggaggaa tggcctgccc ttggggattc agaaattcta aaaaagcagc tgaaacagtg
3481 cagactttta gtcagtgata ttcagacaat tcagcccagt ctaaacagtg tcaatgaagg
3541 tgggcagaag ataaagaatg aagcagagcc agagtttgct tcgagacttg agacagaact
3601 caaagaactt aacactcagt gggatcacat gtgccaacag gtctatgcca gaaaggaggc
```

FIGURE 1 (cont.)

```
3661 cttgaaggga ggtttggaga aaactgtaag cctccagaaa gatctatcag agatgcacga
3721 atggatgaca caagctgaag aagagtatct tgagagagat tttgaatata aaactccaga
3781 tgaattacag aaagcagttg aagagatgaa gagagctaaa gaagaggccc aacaaaaaga
3841 agcgaaagtg aaactcctta ctgagtctgt aaatagtgtc atagctcaag ctccacctgt
3901 agcacaagag gccttaaaaa aggaacttga aactctaacc accaactacc agtggctctg
3961 cactaggctg aatgggaaat gcaagacttt ggaagaagtt tgggcatgtt ggcatgagtt
4021 attgtcatac ttggagaaag caaacaagtg gctaaatgaa gtagaattta aacttaaaac
4081 cactgaaaac attcctggcg gagctgagga aatctctgag gtgctagatt cacttgaaaa
4141 tttgatgcga cattcagagg ataacccaaa tcagattcgc atattggcac agaccctaac
4201 agatggcgga gtcatggatg agctaatcaa tgaggaactt gagacattta attctcgttg
4261 gagggaacta catgaagagg ctgtaaggag gcaaaagttg cttgaacaga gcatccagtc
4321 tgcccaggag actgaaaaat ccttacactt aatccaggag tccctcacat tcattgacaa
4381 gcagttggca gcttatattg cagacaaggt ggacgcagct caaatgcctc aggaagccca
4441 gaaaatccaa tctgatttga caagtcatga gatcagttta gaagaaatga agaaacataa
4501 tcaggggaag gaggctgccc aaagagtcct gtctcagatt gatgttgcac agaaaaaatt
4561 acaagatgtc tccatgaagt ttcgattatt ccagaaacca gccaattttg agctgcgtct
4621 acaagaaagt aagatgattt tagatgaagt gaagatgcac ttgcctgcat tggaaacaaa
4681 gagtgtggaa caggaagtag tacagtcaca gctaaatcat tgtgtgaact tgtataaaag
4741 tctgagtgaa gtgaagtctg aagtggaaat ggtgataaag actggacgtc agattgtaca
4801 gaaaaagcag acggaaaatc ccaaagaact tgatgaaaga gtaacagctt tgaaattgca
4861 ttataatgag ctgggagcaa aggtaacaga aagaaagcaa cagttggaga aatgcttgaa
4921 attgtcccgt aagatgcgaa aggaaatgaa tgtcttgaca gaatggctgg cagctacaga
4981 tatggaattg acaaagagat cagcagttga aggaatgcct agtaatttgg attctgaagt
5041 tgcctgggga aaggctactc aaaaagagat tgagaaacag aaggtgcacc tgaagagtat
5101 cacagaggta ggagaggcct tgaaaacagt tttgggcaag aaggagacgt tggtggaaga
5161 taaactcagt cttctgaata gtaactggat agctgtcacc tcccgagcag aagagtggtt
5221 aaatcttttg ttggaatacc agaaacacat ggaaacttttt gaccagaatg tggaccacat
5281 cacaaagtgg atcattcagg ctgacacact tttggatgaa tcagagaaaa agaaacccca
5341 gcaaaaagaa gacgtgctta agcgtttaaa ggcagaactg aatgacatac gcccaaaggt
5401 ggactctaca cgtgaccaag cagcaaactt gatggcaaac cgcggtgacc actgcaggaa
5461 attagtagag ccccaaatct cagagctcaa ccatcgattt gcagccattt cacacagaat
5521 taagactgga aaggcctcca ttcctttgaa ggaattggag cagtttaact cagatataca
5581 aaaattgctt gaaccactgg aggctgaaat tcagcagggg gtgaatctga aagaggaaga
5641 cttcaataaa gatatgaatg aagacaatga gggtactgta aaagaattgt tgcaaagagg
5701 agacaactta caacaaagaa tcacagatga gagaaagaga gaggaaataa agataaaaca
5761 gcagctgtta cagacaaaac ataatgctct caaggatttg aggtctcaaa gaagaaaaaa
5821 ggctctagaa atttctcatc agtggtatca gtacaagagg caggctgatg atctcctgaa
5881 atgcttggat gacattgaaa aaaaattagc cagcctacct gagcccagag atgaaaggaa
5941 aataaaggaa attgatcggg aattgcagaa gaagaaagag gagctgaatg cagtgcgtag
6001 gcaagctgag ggcttgtctg aggatggggc cgcaatggca gtggagccaa ctcagatcca
6061 gctcagcaag cgctggcggg aaattgagag caaatttgct cagtttcgaa gactcaactt
6121 tgcacaaatt cacactgtcc gtgaagaaac gatgatggtg atgactgaag acatgccttt
6181 ggaaatttct tatgtgcctt ctacttattt gactgaaatc actcatgtct cacaagccct
6241 attagaagtg gaacaacttc tcaatgctcc tgacctctgt gctaaggact ttgaagatct
6301 ctttaagcaa gaggagtctc tgaagaatat aaaagatagt ctacaacaaa gctcaggtcg
6361 gattgacatt attcatagca agaagacagc agcattgcaa agtgcaacgc ctgtggaaag
6421 ggtgaagcta caggaagctc tctcccagct tgatttccaa tgggaaaaag ttaacaaaat
6481 gtacaaggac cgacaagggc gatttgacag atctgttgag aaatggcggc gttttcatta
6541 tgatataaag atatttaatc agtggctaac agaagctgaa cagtttctca gaaagacaca
6601 aattcctgag aattgggaac atgctaaata caaatggtat cttaaggaac tccaggatgg
6661 cattgggcag cggcaaactg ttgtcagaac attgaatgca actggggaag aaataattca
6721 gcaatcctca aaaacagatg ccagtattct acaggaaaaa ttgggaagcc tgaatctgcg
6781 gtggcaggag gtctgcaaac agctgtcaga cagaaaaaag aggctagaag aacaaaagaa
6841 tatcttgtca gaatttcaaa gagatttaaa tgaatttgtt ttatggttgg aggaagcaga
6901 taacattgct agtatcccac ttgaacctgg aaaagagcag caactaaaag aaaagcttga
6961 gcaagtcaag ttactggtgg aagagttgcc cctgcgccag ggaattctca aacaattaaa
7021 tgaaactgga ggacccgtgc ttgtaagtgc tcccataagc ccagaagagc aagataaact
7081 tgaaaataag ctcaagcaga caaatctcca gtggataaag gtttccagag ctttacctga
7141 gaaacaagga gaaattgaag ctcaaataaa agaccttggg cagcttgaaa aaaagcttga
7201 agaccttgaa gagcagttaa atcatctgct gctgtggtta tctcctatta ggaatcagtt
7261 ggaaatttat aaccaaccaa accaagaagg accatttgac gttcaggaaa ctgaaatagc
7321 agttcaagct aaacaaccgg atgtggaaga gattttgtct aaagggcagc atttgtacaa
7381 ggaaaaacca gccactcagc cagtgaagag gaagttagaa gatctgagct ctgagtggaa
7441 ggcggtaaac cgtttacttc aagagctgag ggcaaagcag cctgacctag ctcctggact
```

FIGURE 1 (cont.)

7501 gaccactatt ggagcctctc ctactcagac tgttactctg gtgacacaac ctgtggttac
7561 taaggaaact gccatctcca aactagaaat gccatcttcc ttgatgttgg aggtacctgc
7621 tctggcagat ttcaaccggg cttggacaga acttaccgac tggctttctc tgcttgatca
7681 agttataaaa tcacagaggg tgatggtggg tgaccttgag gatatcaacg agatgatcat
7741 caagcagaag gcaacaatgc aggatttgga acagaggcgt ccccagttgg aagaactcat
7801 taccgctgcc caaaatttga aaaacaagac cagcaatcaa gaggctagaa caatcattac
7861 ggatcgaatt gaaagaattc agaatcagtg ggatgaagta caagaacacc ttcagaaccg
7921 gaggcaacag ttgaatgaaa tgttaaagga ttcaacacaa tggctggaag ctaaggaaga
7981 agctgagcag gtcttaggac aggccagagc caagcttgag tcatggaagg agggtcccta
8041 tacagtagat gcaatccaaa agaaaatcac agaaaccaag cagttggcca aagacctccg
8101 ccagtggcag acaaatgtag atgtggcaaa tgacttggcc ctgaaacttc tccgggatta
8161 ttctgcagat gataccagaa aagtccacat gataacagag aatatcaatg cctcttggag
8221 aagcattcat aaaagggtga gtgagcgaga ggctgctttg gaagaaactc atagattact
8281 gcaacagttc cccctggacc tggaaaagtt tcttgcctgg cttacagaag ctgaaacaac
8341 tgccaatgtc ctacaggatg ctacccgtaa ggaaaggctc ctagaagact ccaagggagt
8401 aaaagagctg atgaaacaat ggcaagacct ccaaggtgaa attgaagctc acacagatgt
8461 ttatcacaac ctggatgaaa acagccaaaa aatcctgaga tccctggaag gttccgatga
8521 tgcagtcctg ttacaaagac gtttggataa catgaacttc aagtggagtg aacttcggaa
8581 aaagtctctc aacattaggt cccatttgga agccagttct gaccagtgga agcgtctgca
8641 cctttctctg caggaacttc tggtgtggct acagctgaaa gatgatgaat taagccggca
8701 ggcacctatt ggaggcgact ttccagcagt tcagaagcag aacgatgtac atagggcctt
8761 caagagggaa ttgaaaacta aagaacctgt aatcatgagt actcttgaga ctgtacgaat
8821 atttctgaca gagcagcctt tggaaggact agagaaactc taccaggagc ccagagagct
8881 gcctcctgag gagagagccc agaatgtcac tcggcttcta cgaaagcagg ctgaggaggt
8941 caatactgag tgggaaaaat tgaacctgca ctccgctgac tggcagagaa aaatagatga
9001 gacccttgaa agactccagg aacttcaaga ggccacggat gagctggacc tcaagctgcg
9061 ccaagctgag gtgatcaagg gatcctggca gcccgtgggc gatctcctca ttgactctct
9121 ccaagatcac ctcgagaaag tcaaggcact tcgaggagaa attgcgcctc tgaaagagaa
9181 cgtgagccac gtcaatgacc ttgctcgcca gcttaccact ttgggcattc agctctcacc
9241 gtataacctc agcactctgg aagacctgaa caccagatgg aagcttctgc aggtggccgt
9301 cgaggaccga gtcaggcagc tgcatgaagc ccacagggac tttggtccag catctcagca
9361 ctttctttcc acgtctgtcc agggtccctg ggagagagcc atctcgccaa acaaagtgcc
9421 ctactatatc aaccacgaga ctcaaacaac ttgctgggac catcccaaaa tgacagagct
9481 ctaccagtct ttagctgacc tgaataatgt cagattctca gcttatagga ctgccatgaa
9541 actccgaaga ctgcagaagg ccctttgctt ggatctcttg agcctgtcag ctgcatgtga
9601 tgccttggac cagcacaacc tcaagcaaaa tgaccagccc atggatatcc tgcagattat
9661 taattgtttg accactattt atgaccgcct ggagcaagag cacaacaatt tggtcaacgt
9721 ccctctctgc gtggatatgt gtctgaactg gctgctgaat gtttatgata cgggacgaac
9781 agggaggatc cgtgtcctgt cttttaaaac tggcatcatt tccctgtgta aagcacattt
9841 ggaagacaag tacagatacc ttttcaagca agtggcaagt tcaacaggat tttgtgacca
9901 gcgcaggctg ggcctccttc tgcatgattc tatccaaatt ccaagacagt tgggtgaagt
9961 tgcatccttt gggggcagta acattgagcc aagtgtccgg agctgcttcc aatttgctaa
10021 taataagcca gagatcgaag cggccctctt cctagactgg atgagactgg aaccccagtc
10081 catggtgtgg ctgcccgtcc tgcacagagt ggctgctgca gaaactgcca agcatcaggc
10141 caaatgtaac atctgcaaag agtgtccaat cattggattc aggtacagga gtctaaagca
10201 ctttaattat gacatctgcc aaagctgctt tttttctggt cgagttgcaa aaggccataa
10261 aatgcactat cccatggtgg aatattgcac tccgactaca tcaggagaag atgttcgaga
10321 ctttgccaag gtactaaaaa acaaatttcg aaccaaaagg tattttgcga agcatccccg
10381 aatgggctac ctgccagtgc agactgtctt agagggggac aacatggaaa ctcccgttac
10441 tctgatcaac ttctggccag tagattctgc gcctgcctcg tcccctcagc tttcacacga
10501 tgatactcat tcacgcattg aacattatgc tagcaggcta gcagaaatgg aaaacagcaa
10561 tggatcttat ctaaatgata gcatctctcc taatgagagc atagatgatg aacatttgtt
10621 aatccagcat tactgccaaa gtttgaacca ggactccccc ctgagccagc ctcgtagtcc
10681 tgcccagatc ttgatttcct tagagagtga ggaaagaggg gagctagaga gaatcctagc
10741 agatcttgag gaagaaaaca ggaatctgca agcagaatat gaccgtctaa agcagcagca
10801 cgaacataaa ggcctgtccc cactgccgtc ccctcctgaa atgatgccca cctctcccca
10861 gagtccccgg gatgctgagc tcattgctga ggccaagcta ctgcgtcaac acaaaggccg
10921 cctggaagcc aggatgcaaa tcctggaaga ccacaataaa cagctggagt cacagttaca
10981 caggctaagg cagctgctgg agcaacccca ggcagaggcc aaagtgaatg gcacaacggt
11041 gtcctctcct tctacctctc tacagaggtc cgacagcagt cagcctatgc tgctccgagt
11101 ggttggcagt caaacttcgg actccatggg tgaggaagat cttctcagtc ctccccagga
11161 cacaagcaca gggttagagg aggtgatgga gcaactcaac aactccttcc ctagttcaag
11221 aggaagaaat accccctggaa agccaatgag agaggacaca atgtaggaag tcttttccac
11281 atggcagatg atttgggcag agcgatggag tccttagtat cagtcatgac agatgaagaa

FIGURE 1 (cont.)

```
11341 ggagcagaat aaatgtttta caactcctga ttcccgcatg gtttttataa tattcataca
11401 acaaagagga ttagacagta agagtttaca agaaataaat ctatattttt gtgaagggta
11461 gtggtattat actgtagatt tcagtagttt ctaagtctgt tattgttttg ttaacaatgg
11521 caggttttac acgtctatgc aattgtacaa aaaagttata agaaaactac atgtaaaatc
11581 ttgatagcta aataacttgc catttcttta tatggaacgc attttgggtt gtttaaaaat
11641 ttataacagt tataaagaaa gattgtaaac taaagtgtgc tttataaaaa aaagttgttt
11701 ataaaaaccc ctaaaaacaa aacaaacaca cacacacaca catacacaca cacacacaaa
11761 actttgaggc agcgcattgt tttgcatcct tttggcgtga tatccatatg aaattcatgg
11821 ctttttcttt ttttgcatat taaagataag acttcctcta ccaccacacc aaatgactac
11881 tacacactgc tcatttgaga actgtcagct gagtggggca ggcttgagtt ttcatttcat
11941 atatctatat gtctataagt atataaatac tatagttata tagataaaga gatacgaatt
12001 tctatagact gacttttttcc atttttttaaa tgttcatgtc acatcctaat agaaagaaat
12061 tacttctagt cagtcatcca ggcttacctg cttggtctag aatggatttt tcccggagcc
12121 ggaagccagg aggaaactac accacactaa aacattgtct acagctccag atgtttctca
12181 ttttaaacaa ctttccactg acaacgaaag taaagtaaag tattggattt ttttaaaggg
12241 aacatgtgaa tgaatacaca ggacttatta tatcagagtg agtaatcggt tggttggttg
12301 attgattgat tgattgatac attcagcttc ctgctgctag caatgccacg atttagattt
12361 aatgatgctt cagtggaaat caatcagaag gtattctgac cttgtgaaca tcagaaggta
12421 ttttttaact cccaagcagt agcaggacga tgatagggct ggagggctat ggattcccag
12481 cccatccctg tgaaggagta ggccactctt taagtgaagg attggatgat tgttcataat
12541 acataaagtt ctctgtaatt acaactaaat tattatgccc tcttctcaca gtcaaaagga
12601 actgggtggt ttggtttttg ttgctttttt agatttattg tcccatgtgg gatgagtttt
12661 taaatgccac aagacataat ttaaaataaa taaactttgg gaaaaggtgt aagacagtag
12721 ccccatcaca tttgtgatac tgacaggtat caacccagaa gcccatgaac tgtgtttcca
12781 tcctttgcat ttctctgcga gtagttccac acaggtttgt aagtaagtaa gaaagaaggc
12841 aaattgattc aaatgttaca aaaaaaccct tcttggtgga ttagacaggt taaatatata
12901 aacaaacaaa caaaaattgc tcaaaaaaga ggagaaaagc tcaagaggaa aagctaagga
12961 ctggtaggaa aaagctttac tctttcatgc cattttattt ctttttgatt tttaaatcat
13021 tcattcaata gataccaccg tgtgacctat aattttgcaa atctgttacc tctgacatca
13081 agtgtaatta gcttttggag agtgggctga catcaagtgt aattagcttt tggagagtgg
13141 gttttgtcca ttattaataa ttaattaatt aacatcaaac acggcttctc atgctatttc
13201 tacctcactt tggttttggg gtgttcctga taattgtgca cacctgagtt cacagcttca
13261 ccacttgtcc attgcgttat tttctttttc ctttataatt ctttctttt ccttcataat
13321 tttcaaaaga aaacccaaag ctctaaggta acaaattacc aaattacatg aagatttggt
13381 ttttgtcttg catttttttc ctttatgtga cgctggacct tttctttacc caaggatttt
13441 taaaactcag atttaaaaca aggggttact ttacatccta ctaagaagtt taagtaagta
13501 agtttcattc taaaatcaga ggtaaataga gtgcataaat aattttgttt taatcttttt
13561 gtttttcttt tagacacatt agctctggag tgagtctgtc ataatatttg aacaaaaatt
13621 gagagcttta ttgctgcatt ttaagcataa ttaatttgga cattatttcg tgttgtgttc
13681 tttataacca ccgagtatta aactgtaaat cataatgtaa ctgaagcata aacatcacat
13741 ggcatgtttt gtcattgttt tcaggtactg agttcttact tgagtatcat aatatattgt
13801 gttttaacac caacactgta acatttacga attatttttt taaacttcag ttttactgca
13861 ttttcacaac atatcagact tcaccaaata tatgccttac tattgtatta tagtactgct
13921 ttactgtgta tctcaataaa gcacgcagtt atgttac
```

FIGURE 2 (Mouse Dystrophin cDNA, Acc. No. M68859, SEQ ID NO:2)

```
   1 cctcactcac ttgcccctta caggactcag ctcttgaagg caatagcttt atagaaaaaa
  61 cgaataggaa gacttgaagt gctatttttt ttttttttt tgtcaaggct gctgaagttt
 121 attggcttct catcgtacct aagcctcctg gagcaataaa actgggagaa acttttacca
 181 agatttttat ccctgccttg atatatactt tttcttccaa atgctttggt gggaagaagt
 241 agaggactgt tatgaaagag aagatgttca aaagaaaaca ttcacaaaat ggataaatgc
 301 acaatttttct aagtttggaa agcaacacat agacaacctc ttcagtgacc tgcaggatgg
 361 aaaacgcctc ctagacctct tggaaggcct tacagggcaa aaactgccaa aagaaaaggg
 421 atctacaaga gttcatgccc tgaacaatgt caacaaggca ctgcgggtct tacagaaaaa
 481 taatgttgat ttagtgaata taggaagcac tgacatagtg gatggaaatc ataaactcac
 541 tcttggtttg atttggaata taatcctcca ctggcaggtc aaaaatgtga tgaaaactat
 601 catggctgga ttgcagcaaa ccaacagtga aaagattctt ctgagctggg ttcgacagtc
 661 aacacgtaat tatccacagg ttaacgtcat caacttcacc tctagctggt ccgacgggtt
 721 ggctttgaat gctcttatcc atagtcacag gcccgacctg tttgattgga atagtgtggt
 781 ttcacagcac tcagccaccc aaagactgga acatgccttc aacattgcaa aatgccagtt
 841 aggcatagaa aaacttcttg atcctgaaga tgttgctacc acttatccag acaagaagtc
 901 catcttaatg tacatcacat cactctttca agttttgcca caacaagtga gcattgaagc
 961 cattcaagaa gtggaaatgt tgcccaggac atcttcaaaa gtaactagag aagaacattt
1021 tcaattacat caccagatgc attactctca acagatcaca gtcagtctag cacagggcta
1081 tgaacaaact tcttcatctc ctaagcctcg attcaagagt tatgccttca cacaggctgc
1141 ttatgttgcc acctctgatt ccacacagag cccctatcct tcacagcatt tggaagctcc
1201 cagagacaag tcacttgaca gttcattgat ggagacggaa gtaaatctgg atagttacca
1261 aactgcttta gaagaagtac tttcatggct tctttctgcc gaggatacat tgcgagcaca
1321 aggagagatt tcaaatgatg ttgaagaagt gaaagaacag tttcatgctc atgagggatt
1381 catgatggat ctgacatctc atcaaggact tgttggtaat gttctacagt taggaagtca
1441 actagttgga aaagggaaat tatcagaaga tgaagaagct gaagtgcaag aacaaatgaa
1501 tctcctaaat tcaagatggg aatgtctcag ggtagctagc atggaaaaac aaagcaaatt
1561 acacaaagtt ctaatggatc tccagaatca gaaattaaaa gaactagatg actggttaac
1621 aaaaactgaa gagagaacta agaaaatgga ggaagagccc tttggacctg atcttgaaga
1681 tctaaaatgc caagtacaac aacataaggt gcttcaagaa gatctagaac aggagcaggt
1741 cagggtcaac tcgctcactc acatggtagt agtggttgat gaatccagcg gtgatcatgc
1801 aacagctgct ttggaagaac aacttaaggt actgggagat cgatgggcaa atatctgcag
1861 atggactgaa gaccgctgga ttgttttaca agatattctt ctaaaatggc agcattttac
1921 tgaagaacag tgccttttta gtacatggct ttcagaaaaa gaagatgcaa tgaagaacat
1981 tcagacaagt ggctttaaag atcaaaatga aatgatgtca agtcttcaca aaatatctac
2041 tttaaaaata gatctagaaa agaaaaagcc aaccatggaa aaactaagtt cactcaatca
2101 agatctactt tcggcactga aaaataagtc agtgactcaa aagatggaaa tctggatgga
2161 aaactttgca caacgttggg acaatttaac ccaaaaactt gaaaagagtt cagcacaaat
2221 ttcacaggct gtcaccacca ctcaaccatc cctaacacag acaactgtaa tggaaacggt
2281 aactatggtg accacaaggg aacaaatcat ggtaaaacat gcccaagagg aacttccacc
2341 accacctcct caaaagaaga ggcagataac tgtggattct gaactcagga aaaggttgga
2401 tgtcgatata actgaacttc acagttggat tactcgttca gaagctgtat tacagagttc
2461 tgaatttgca gtctatcgaa aagaaggcaa catctcagac ttgcaagaaa aagtcaatgc
2521 catagcacga gaaaaagcag agaagttcag aaaactgcaa gatgccagca gatcagctca
2581 ggccctggtg gaacagatgg caaatgaggg tgttaatgct gaaagtatca gacaagcttc
2641 agaacaactg aacagccggt ggacagaatt ctgccaattg ctgagtgaga gagttaactg
2701 gctagagtat caaaccaaca tcattacctt ttataatcag ctacaacaat tggaacagat
2761 gacaactact gccgaaaact tgttgaaaac ccagtctacc accctatcag agccaacagc
2821 aattaaaagc eagttaaaaa tttgtaagga tgaagtcaac agattgtcag ctcttcagcc
2881 tcaaattgag caattaaaaa ttcagagtct acaactgaaa gaaaagggac aggggccaat
2941 gtttctggat gcagactttg tggcctttac taatcatttt aaccacatct ttgatggtgt
3001 gagggccaaa gagaaagagc tacagacaat ttttgacact ttaccaccaa tgcgctatca
3061 ggagacaatg agtagcatca ggacgtggat ccagcagtca gaaagcaaac tctctgtacc
3121 ttatcttagt gttactgaat atgaaataat ggaggagaga ctcgggaaat tacaggctct
3181 gcaaagttct ttgaaagagc aacaaaatgg cttcaactat ctgagtgaca ctgtgaagga
3241 gatggccaag aaagcacctt cagaaatatg ccagaaatat ctgtcagaat ttgaagagat
3301 tgaggggcac tggaagaaac tttcctccca gttggtggaa agctgccaaa agctagaaga
3361 acatatgaat aaacttcgaa aatttcagaa tcacataaaa accttacaga aatggatggc
3421 tgaagttgat gttttcctga aagaggaatg gcctgccctg ggggatgctg aaatcctgaa
3481 aaaacagctc aaacaatgca gacttttagt tggtgatatt caaacaattc agcccagttt
3541 aaatagtgtt aatgaaggtg ggcagaagat aaagagtgaa gctgaacttg agtttgcatc
3601 cagactggag acagaactta gagagcttaa cactcagtgg gatcacatat gccgccaggt
```

FIGURE 2 (cont.)

```
3661 ctacaccaga aaggaagcct taaaggcagg tttggataaa accgtaagcc tccaaaaaga
3721 tctatcagag atgcatgagt ggatgacaca agctgaagaa gaatatctag agagagattt
3781 tgaatataaa actccagatg aattacagac tgctgttgaa gaaatgaaga gagctaaaga
3841 agaggcacta caaaaagaaa ctaaagtgaa actccttact gagactgtaa atagtgtaat
3901 agctcacgct ccaccctcag cacaagaggc cttaaaaaag gaacttgaaa ctctgaccac
3961 caactaccaa tggctgtgca ccaggctgaa tggaaaatgc aaaactttgg aagaagtttg
4021 ggcatgttgg catgagttat tgtcatattt agagaaagca aacaagtggc tcaatgaagt
4081 agaattgaaa cttaaaacca tggaaaatgt tcctgcagga cctgaggaaa tcactgaagt
4141 gctagaatct cttgaaaatc tgatgcatca ttcagaggag aacccaaatc agattcgtct
4201 attggcacag actcttacag atggaggagt catggatgaa ctgatcaatg aggagcttga
4261 gacgtttaat tctcgttgga gggaactaca tgaagaggct gtgaggaaac aaaagttgct
4321 tgaacagagt atccagtctg cccaggaaat tgaaaagtcc ttgcacttaa ttcaggagtc
4381 gcttgaattc attgacaagc agttggcagc ttatatcact gacaaggtgg atgcagctca
4441 aatgcctcag gaagcccaga aaatccaatc agatttgaca agtcatgaga taagtttaga
4501 agaaatgaag aaacataacc aggggaagga tgccaaccaa agggttcttt cacaaattga
4561 tgttgcacag aaaaaattac aagatgtctc catgaaattt cgattattcc aaaaaccagc
4621 caattttgaa caacgtctag aggaaagtaa gatgatttta gatgaagtca agatgcattt
4681 gcctgcattg gaaaccaaga gtgttgaaca ggaagtaatt cagtcacaac taagtcattg
4741 tgtgaacttg tataaaagcc tgagtgaagt caagtctgaa gtggaaatgg tgattaaaac
4801 cggacgtcaa attgtacaga aaaagcagac agaaaatccc aaagagcttg atgaacgagt
4861 aacagctttg aaattgcatt acaatgagtt gggtgcgaag gtaacagaga gaaagcaaca
4921 gttggagaaa tgcttgaagt tgtcccgtaa gatgagaaag gaaatgaatg tcttaacaga
4981 atggctggca gcaacagata cagaattgac gaagagatca gcagttgaag gaatgccaag
5041 taatttggat tctgaagttg cctggggaaa ggctactcaa aaagagattg agaaacagaa
5101 ggctcacttg aagagtgtta cagaattagg agagtctttg aaaatggtgt tgggcaagaa
5161 agaaaccttg gtagaagata aactgagtct tctgaacagt aactggatag ctgtcacctc
5221 cagagtagaa gaatggctaa atcttttgtt ggaataccag aaacacatgg aaacctttga
5281 tcagaacata gaacaaatca caaagtggat cattcatgca gatgaacttt tagatgagtc
5341 tgaaaagaag aaaccacaac aaaaggaaga cattcttaag cgtttaaagg ctgaaatgaa
5401 tgacatgcgc ccaaaggtgg actccacacg tgaccaagca gcaaaattga tggcaaaccg
5461 cggtgaccac tgcaggaaag tagtagagcc ccaaatctct gagctcaacc gtcgatttgc
5521 agctatttct cacagaatta agactggaaa ggcctccatt cctttgaagg aattggagca
5581 gtttaactca gatatacaaa aattgcttga accactggag gctgaaattc agcagggggt
5641 gaatctgaaa gaggaagact tcaataaaga tatgagtgaa gacaatgagg gtactgtaaa
5701 tgaattgttg caaagaggag acaacttaca acaaagaatc acagatgaga gaaagcgaga
5761 ggaaataaag ataaaacagc agctgttaca gacaaaacat aatgctctca aggatttgag
5821 gtctcaaaga agaaaaaagg ccctagaaat ttctcaccag tggtatcagt acaagaggca
5881 ggctgatgat ctcctgaaat gcttggatga aattgaaaaa aaattagcca gcctacctga
5941 acccagagat gaaagaaaat taaaggaaat tgatcgtgaa ttgcagaaga agaaagagga
6001 gctgaatgca gtgcgcaggc aagctgaggg cttgtctgag aatggggccg caatggcagt
6061 ggagccaact cagatccagc tcagcaagcg ctggcggcaa attgagagca attttgctca
6121 gtttcgaaga ctcaactttg cacaaattca cactctccat gaagaaacta tggtagtgac
6181 gactgaagat atgcctttgg atgtttctta tgtgccttct acttatttga ccgagatcag
6241 tcatatctta caagctcttt cagaagttga tcatcttcta aatactcctg aactctgtgc
6301 taaagatttt gaagatcttt ttaagcaaga ggagtctctt aagaatataa aagacaattt
6361 gcaacaaatc tcaggtcgga ttgatattat tcacaagaag aagacagcag ccttgcaaag
6421 tgccacctcc atggaaaagg tgaaagtaca ggaagccgtg gcacagatgg atttccaggg
6481 ggaaaaactt catagaatgt acaaggaacg acaagggcga ttcgacagat cagttgaaaa
6541 atggcgacac tttcattatg atatgaaggt atttaatcaa tggctgaatg aagttgaaca
6601 gtttttcaaa aagacacaaa atcctgaaaa ctgggaacat gctaaataca aatggtatct
6661 taaggaactc caggatggca ttgggcagcg tcaagctgtt gtcagaacac tgaatgcaac
6721 tggggaagaa ataattcaac agtcttcaaa aacagatgtc aatattctac aagaaaaatt
6781 aggaagcttg agtctgcggt ggcacgacat ctgcaaagag ctggcagaaa ggagaaagag
6841 gattgaagaa caaaagaatg tcttgtcaga atttcaaaga gatttaaatg aatttgtttt
6901 gtggctggaa gaagcagata acattgctat tactccactt ggagatgagc agcagctaaa
6961 agaacaactt gaacaagtca agttactggc agaagagttg cccctgcgcc agggaattct
7021 aaaacaatta aatgaaacag gaggagcagt acttgtaagt gctcccataa ggccagaaga
7081 gcaagataaa cttgaaaaga agctcaaaca gacaaatctc cagtggataa aggtctccag
7141 agctttacct gagaaacaag gagagcttga ggttcactta aaagatttta ggcagcttga
7201 agagcagctg gatcacctgc ttctgtggct ctctcctatt agaaaccagt tggaaattta
7261 taaccaacca agtcaggcag gaccgtttga cataaaggag attgaagtaa cagttcacgg
7321 taaacaagcg gatgtggaaa ggcttttgtc gaaagggcag catttgtata aggaaaaacc
7381 aagcactcag ccagtgaaga ggaagttaga agatctgagg tctgagtggg aggctgtaaa
7441 ccatttactt cgggagctga ggacaaagca gcctgaccgt gcccctggac tgagcactac
```

FIGURE 2 (cont.)

```
7501 tggagcctct gccagtcaga ctgttactct agtgacacaa tctgtggtta ctaaggaaac
7561 tgtcatctcc aaactagaaa tgccatcttc tttgctgttg gaggtacctg cactggcaga
7621 cttcaaccga gcttggacag aacttacaga ctggctgtct ctgcttgatc gagttataaa
7681 atcacagaga gtgatggtgg gtgatctgga agacatcaat gaaatgatca tcaaacagaa
7741 ggcaacactg caagatttgg aacagagacg cccccaattg gaagaactca ttactgctgc
7801 ccagaatttg aaaaacaaaa ccagcaatca agaagctaga acaatcatta ctgatcgaat
7861 tgaaagaatt cagattcagt gggatgaggt tcaagaacag ctgcagaaca ggagacaaca
7921 gttgaatgaa atgttaaagg attcaacaca atggctggaa gctaaggaag aagccgaaca
7981 ggtcatagga caggtcagag gcaagcttga ctcatggaaa gaaggtcctc acacagtaga
8041 tgcaatccaa aagaagatca cagaaaccaa gcagttggcc aaagacctcc gtcaacggca
8101 gataagtgta gacgtggcaa atgatttggc actgaaactt cttcgggact attctgctga
8161 tgataccaga aaagtacaca tgataacaga gaatatcaat acttcttggg gaaacattca
8221 taaaagagta agtgagcaag aggctgcttt ggaagaaact catagattac tgcagcagtt
8281 ccctctggac ctggagaagt ttctttcctg gattacggaa gcagaaacaa ctgccaatgt
8341 cctacaggac gcttcccgta aggagaagct cctagaagac tccaggggag tcagagagct
8401 gatgaaacca tggcaagatc tccaaggaga aattgaaact cacacagata tctatcacaa
8461 tcttgatgaa aatggccaaa aaatcctgag atccctggaa ggttcggatg aagcacccct
8521 gttacaaaga cgtttggata acatgaattt caagtggagt gaacttcaga aaaagtctct
8581 caacattagg tcccatttgg aagcaagttc tgaccagtgg aagcgtttgc atctttctct
8641 tcaggaactt cttgtttggc tacagctgaa agatgatgaa ctgagccgtc aggcacccat
8701 cggtggtgat ttcccagcag ttcagaagca gaatgatata catagggcct tcaagaggga
8761 attgaaaact aaagaacctg taatcatgag tactctggag actgtgagaa tatttctgac
8821 agagcagcct ttggaaggac tagagaaact ctaccaggag cccagagaac tgcctcctga
8881 agaaagagct cagaatgtca ctcggctcct acgaaagcag gctgaagagg tcaacgctga
8941 atgggacaaa ttgaacctgc gctcagctga ttggcagaga aaaatagatg aagctcttga
9001 aagactccag gaacttcagg aagctgccga tgaactggac ctcaagttgc gccaagctga
9061 ggtgatcaag ggatcctggc agccagtggg ggatctcctc attgactctc tgcaagatca
9121 ccttgaaaaa gtcaaggcac ttcggggaga aattgcacct cttaaagaga atgtcaatcg
9181 tgtcaatgac cttgcacatc agctgaccac actgggcatt cagctctcac cttataacct
9241 cagcactttg gaagatctga ataccagatg gaggcttcta caggtggctg tggaggaccg
9301 tgtcagacag ctgcatgaag cccacaggga ctttggtcct gcatcccagc acttcctttc
9361 cacttcagtt cagggtccct gggagagagc catctcacca aacaaagtgc cctactatat
9421 caaccacgag acccaaacca cttgttggga ccaccccaaa atgacagagc tctaccagtc
9481 tttagctgac ctgaataatg tcaggttctc cgcgtatagg actgccatga agctcagaag
9541 gctccagaag gccctttgct tggatctctt gagcctgtca gctgcatgtg atgccctgga
9601 ccagcacaac ctcaagcaaa atgaccagcc catggatatc ctgcagataa ttaactgttt
9661 gactacaatt tatgatcgtc tggagcaaga gcacaacaat ctggtcaatg tccctctctg
9721 tgtggatatg tgtctcaact ggcttctcaa tgtttatgat acgggacgaa cagggaggat
9781 ccgtgtcctg tcttttaaaa ctggcatcat ttctctgtgt aaagcacact tggaagacaa
9841 gtacagatac cttttcaagc aagtggcaag ttcaactggc ttttgtgacc agcgtaggct
9901 gggtcttctt ctgcatgatt ctattcaaat cccaagacag ttgggtgaag ttgcttcctt
9961 tgggggcagt aacattgagc cgagtgtcag gagctgcttc caatttgcca ataataaacc
10021 tgagattgaa gctgctctct tccttgactg gatgcgcctg gaaccccagt ctatggtgtg
10081 gctgcccgtc ttgcacagag tggctgctgc tgaaactgcc aagcatcaag ccaagtgtaa
10141 catctgtaag gagtgtccaa tcattggatt caggtacaga agcctaaagc attttaatta
10201 tgacatctgc caaagttgct ttttttctgg ccgagttgca aagggccata aaatgcacta
10261 ccccatggta gagtattgca ctccgactac atccggagaa gatgttcgcg acttcgccaa
10321 ggtactaaaa aacaaatttc gaaccaaaag gtattttgcg aagcatcccc gaatgggcta
10381 cctgccagtg cagactgtgt tagaggggga caacatggaa actcccgtta ctctgatcaa
10441 cttctggcca gtagattctg cgcctgcctc gtcccccag ctttcacacg atgatactca
10501 ttcacgcatt gaacattatg ctagcaggct agcagaaatg gaaaacagca atggatctta
10561 tctaaatgat agcatctctc ctaatgagag catagatgat gaacatttgt taatccagca
10621 ttactgccaa agtttgaacc aggactcccc cctgagccag cctcgtagtc ctgcccagat
10681 cttgatttcc ttagagagtg aggaaagagg ggagctagag agaatcctag cagatcttga
10741 ggaagaaaac aggaatctgc aagcagaata tgatcgcctg aagcagcagc atgagcataa
10801 aggcctgtct ccactgccat ctcctcctga gatgatgccc acctctcctc agagtcccag
10861 ggatgctgag ctcattgctg aggctaagct actgcgccaa cacaaaggac gcctggaagc
10921 caggatgcaa atcctggaag accacaataa acagctggag tctcagttac atagactgag
10981 acagctcctg gagcagcccc aggctgaagc taaggtgaat ggcaccacgg tgtcctctcc
11041 ttccacctct ctgcagaggt cagatagcag tcagcctatg ctgctccgag tggttggcag
11101 tcaaacttca gaatctatgg gtgaggaaga tcttctgagt cctccccagg acacaagcac
11161 agggttagaa gaagtgatgg agcaactcaa caactccttc cctagttcaa gaggaagaaa
11221 tgcccccgga aagccaatga gagaggacac aatgtaggaa gcctttcca catggcagat
11281 gatttgggca gagcgatgga gtccttagtt tcagtcatga cagatgaaga aggagcagaa
```

FIGURE 2 (cont.)

```
11341 taaatgtttt acaactcctg attcccgcat ggtttttata atattcgtac aacaaagagg
11401 attagacagt aagagtttac aagaaataaa atctatattt ttgtgaaggg tagtggtact
11461 atactgtaga tttcagtagt ttctaagtct gttattgttt tgttaacaat ggcaggtttt
11521 acacgtctat gcaattgtac aaaaaagtta aaagaaaaca tgtaaaatct tgatagctaa
11581 ataacttgcc atttctttat atggaacgca ttttgggttg tttaaaaatt tataacagtt
11641 ataaagagag attgtaaact aaagtgtgct ttataaaaaa agttgtttat aaaaacccct
11701 aaacaaacac acacgcacac acacacacac acacacacac acacacacac gcacacatac
11761 atgcacgaac ccaccacaca cacacacaca cacacacaca ctgaggcagc acattgtttt
11821 gcattacttt agcgtggtat tcatatggaa ttcatgacgt ttttttattt tcttgcatac
11881 gaaccccacc aaatgactgc ttcatattgc tcttttgaga attgttgact gagtggggct
11941 ggctatgggc tttcatttta tacatctata tgtctacaag tatataaata ctataggtat
12001 atagataaat agatatgaag ttacttcttc aaatgttctt gccacttcct aatggaaatt
12061 gcttctagtc atctgggctt atctgcttgg gcaagagtga attttccctg gagcccaaag
12121 ccaggagact accgccacac taaaatattg tctagggctc cagatgtttc tagttttaaa
12181 ctttccactg agagctagag gattcatttt tttcaaggaa catgcgaatg aatacacagg
12241 acttactatc atagtaattt gttggctgat atattcaact tcctactgtt gggttatatt
12301 taatgatgtt tctgcaatag aacatcagat gacatttta actcccagac agtaggagga
12361 agatggtagg agctaaaggt tgcggctcct cagtcaattt atatgagggg agcaacaact
12421 ctgtaaaaga atggatgaat atttacaact atacatataa acatctctat aattacaact
12481 aaattgttct gccctcttca taaactcaac ctgaagtggg tggttttgtt gttgttgttg
12541 ttgttgttgt tgatgatgat gatgaatttt agattttaga ttttttgggt tttttttttct
12601 tcattgtgat gatttttttt tttaatgctg caagacttag gattactgtt aagaaagtaa
12661 cccaatcaca ttgtgaccct ggtgaatatc agtccagaag cccatgaact gcatttgtct
12721 cctttgcatt ggtttccctg caagtaactc cacacaggat tgtgggtgag aaggcacagt
12781 ggttggaaag ttttgagagc aaaagcgtct ccaaactctc tggtctagtt gacgggctga
12841 aatgtctaaa caaatgcaag tcattgaacc aggagaaaaa gtgcaacaga aagctaagga
12901 ctgctaggaa gagctttact cctctcatgc cagtttcttc ttcttagcat ttaaagagca
12961 ttctctcaat agaaatcact gtcctatcat tttgcaaatc tgttacctct aacgtcaagt
13021 gtaattaact tctagcgagt gggttttgtc cattattaat tgtaattaac atcaaacaca
13081 gcttctcatg ctatttctac ctcactttgg ttttggggtg tttctagtaa ttgtgcacac
13141 ctaatttcac aacttcacca cttgtctgtt gtgtggacac cagtttcctt ttttcattta
13201 taatttccaa aagaaaaccc aaagctctaa gataacaaat tgaaatttgg ttctggtctt
13261 gcttttctct ctctctctcc tttatgtggc actgggcatt ttctttatcc aaggatttgt
13321 tttcaccaag atttaaaaca aggggttcct ttcctactaa gaagttttaa gtttcattct
13381 aaaatccaag gtagatagag tgcatagttt tgttttaatc ttttcgtttt atcttttaga
13441 tattagttct ggagtgaatc tatcaaaata tttgaataaa aactgagagc tttattgctg
13501 attttaagca taatttggac atcatttcat gttctttata accatcaagt attaaagtgt
13561 aaatcataat cagtgtaact gaagcataat catcacatgg catgtatcat cattgtctcc
13621 aggtactgga ctcttacttg agtatcataa tagattgtgt tttaacacca acactgtaac
13681 atttactaat tatttttta aacttcagtt ttactgcatt ttcacaacat atcagatttc
13741 accaaatata tgccttacta ttgtattata ttactgcttt actgtgtatc tcaataaagc
13801 acgcagttat gttac
```

FIGURE 3 (Human Utrophin cDNA, Acc. No. X69086, SEQ ID NO:3)

```
   1 atggccaagt atggagaaca tgaagccagt cctgacaatg ggcagaacga attcagtgat
  61 atcattaagt ccagatctga tgaacacaat gacgtacaga agaaaacctt taccaaatgg
 121 ataaatgctc gattttcaaa gagtgggaaa ccacccatca atgatatgtt cacagacctc
 181 aaagatggaa ggaagctatt ggatcttcta gaaggcctca caggaacatc actgccaaag
 241 gaacgtggtt ccacaagggt acatgcctta aataacgtca acagagtgct gcaggtttta
 301 catcagaaca atgtggaatt agtgaatata gggggaactg acattgtgga tggaaatcac
 361 aaactgactt tggggttact ttggagcatc attttgcact ggcaggtgaa agatgtcatg
 421 aaggatgtca tgtcggacct gcagcagacg aacagtgaga agatcctgct cagctgggtg
 481 cgtcagacca ccaggcccta cagccaagtc aacgtcctca acttcaccac cagctggaca
 541 gatggactcg cctttaatgc tgtcctccac cgacataaac ctgatctctt cagctgggat
 601 aaagttgtca aaatgtcacc aattgagaga cttgaacatg ccttcagcaa ggctcaaact
 661 tatttgggaa ttgaaaagct gttagatcct gaagatgttg ccgttcggct tcctgacaag
 721 aaatccataa ttatgtattt aacatctttg tttgaggtgc tacctcagca agtcaccata
 781 gacgccatcc gtgaggtaga gacactccca aggaaatata aaaaagaatg tgaagaagag
 841 gcaattaata tacagagtac agcgcctgag gaggagcatg agagtccccg agctgaaact
 901 cccagcactg tcactgaggt cgacatggat ctggacagct atcagattgc gttggaggaa
 961 gtgctgacct ggttgctttc tgctgaggac actttccagg agcaggatga tatttctgat
1021 gatgttgaag aagtcaaaga ccagtttgca acccatgaag cttttatgat ggaactgact
1081 gcacaccaga gcagtgtggg cagcgtcctg caggcaggca accaactgat aacacaagga
1141 actctgtcag acgaagaaga atttgagatt caggaacaga tgaccctgct gaatgctaga
1201 tgggaggctc ttagggtgga gagtatggac agacagtccc ggctgcacga tgtgctgatg
1261 gaactgcaga agaagcaact gcagcagctc tccgcctggt taacactcac agaggagcgc
1321 attcagaaga tggaaacttg ccccctggat gatgatgtaa aatctctaca aaagctgcta
1381 gaagaacata aaagtttgca aagtgatctt gaggctgaac aggtgaaagt aaattcacta
1441 actcacatgg tggtcattgt tgatgaaaac agtggtgaga gcgctacagc tatcctagaa
1501 gaccagttac agaaacttgg tgagcgctgg acagcagtat gccgttggac tgaagaacgc
1561 tggaataggt tacaagaaat caatatattg tggcaggaat tattggaaga acagtgcttg
1621 ttgaaagctt ggttaaccga aaaagaagag gctttaaata aagtccagac aagcaacttc
1681 aaagaccaaa aggaactaag tgtcagtgtt cgacgtctgg ctattttgaa ggaagacatg
1741 gaaatgaagc gtcaaacatt ggatcagctg agtgagattg gccaggatgt gggacaatta
1801 cttgataatt ccaaggcatc taagaagatc aacagtgact cagaggaact gactcaaaga
1861 tgggattctt tggttcagag actagaagat tcctccaacc aggtgactca ggctgtagca
1921 aagctgggga tgtctcagat tcctcagaag gacctttttgg agactgttcg tgtaagagaa
1981 caagcaatta caaaaaaatc taagcaggaa ctgcctcctc ctcctccccc aaagaagaga
2041 cagatccatg tggatattga agctaagaaa aagtttgatg ctataagtgc agagctgttg
2101 aactggattt tgaaatggaa aactgccatt cagaccacag agataaaaga gtatatgaag
2161 atgcaagaca cttccgaaat gaaaaagaag ttgaaggcat tagaaaaaga acagagagaa
2221 agaatcccca gagcagatga attaaaccaa actggacaaa tccttgtgga gcaaatggga
2281 aaagaaggcc ttcctactga agaaataaaa aatgttctgg agaaggtttc atcagaatgg
2341 aagaatgtat ctcaacattt ggaagatcta gaaagaaaga ttcagctaca ggaagatata
2401 aatgcttatt tcaagcagct tgatgagctt gaaaaggtca tcaagacaaa ggaggagtgg
2461 gtaaaacaca cttccatttc tgaatcttcc cggcagtcct tgccaagctt gaaggattcc
2521 tgtcagcggg aattgacaaa tcttcttggc cttcacccca aaattgaaat ggctcgtgca
2581 agctgctcgg ccctgatgtc tcagccttct gccccagatt ttgtccagcg gggcttcgat
2641 agctttctgg gccgctacca agctgtacaa gaggctgtag aggatcgtca acaacatcta
2701 gagaatgaac tgaagggcca acctggacat gcatatctgg aaacattgaa aacactgaaa
2761 gatgtgctaa atgattcaga aaataaggcc caggtgtctc tgaatgtcct taatgatctt
2821 gccaaggtgg agaaggccct gcaagaaaaa aagacccttg atgaaatcct tgagaatcag
2881 aaacctgcat tacataaact tgcagaagaa acaaaggctc tggagaaaaa tgttcatcct
2941 gatgtagaaa aattatataa gcaagaattt gatgatgtgc aaggaaagtg gaacaagcta
3001 aaggtcttgg tttccaaaga tctacatttg cttgaggaaa ttgctctcac actcagagct
3061 tttgaggccg attcaacagt cattgagaag tggatggatg gcgtgaaaga cttcttaatg
3121 aaacagcagg ctgcccaagg agacgacgca ggtctacaga ggcagttaga ccagtgctct
3181 gcatttgtta atgaaataga aacaattgaa tcatctctga aaaacatgaa ggaaatagag
3241 actaatcttc gaagtggtcc agttgctgga ataaaaactt gggtgcagac aagactaggt
3301 gactaccaaa ctcaactgga gaaacttagc aaggagatcg ctactcaaaa aagtaggttg
3361 tctgaaagtc aagaaaaagc tgcgaacctg aagaaagact tggcagagat gcaggaatgg
3421 atgacccagg ccgaggaaga atatttggag cgggattttg agtacaagtc accagaagag
3481 cttgagagtg ctgtggaaga gatgaagagg gcaaaagagg atgtgttgca gaaggaggtg
3541 agagtgaaga ttctcaagga caacatcaag ttattagctg ccaaggtgcc ctctggtggc
3601 caggagttga cgtctgagct gaatgttgtg ctggagaatt accaacttct ttgtaataga
```

FIGURE 3 (cont.)

3661 attcgaggaa agtgccacac gctagaggag gtctggtctt gttggattga actgcttcac
3721 tatttggatc ttgaaactac ctggttaaac actttggaag agcggatgaa gagcacagag
3781 gtcctgcctg agaagacgga tgctgtcaac gaagccctgg agtctctgga atctgttctg
3841 cgccacccgg cagataatcg cacccagatt cgagagcttg gccagactct gattgatggg
3901 gggatcctgg atgatataat cagtgagaaa ctggaggctt tcaacagccg atatgaagat
3961 ctaagtcacc tggcagagag caagcagatt tctttggaaa agcaactcca ggtgctgcgg
4021 gaaactgacc agatgcttca agtcttgcaa gagagcttgg gggagctgga caaacagctc
4081 accacatacc tgactgacag gatagatgct ttccaagttc cacaggaagc tcagaaaatc
4141 caagcagaga tctcagccca tgagctaacc ctagaggagt tgagaagaaa tatgcgttct
4201 cagcccctga cctccccaga gagtaggact gccagaggag gaagtcagat ggatgtgcta
4261 cagaggaaac tccgagaggt gtccacaaag ttccagcttt tccagaagcc agctaacttc
4321 gagcagcgca tgctggactg caagcgtgtg ctggatggcg tgaaagcaga acttcacgtt
4381 ctggatgtga aggacgtaga ccctgacgtc atacagacgc acctggacaa gtgtatgaaa
4441 ctgtataaaa ctttgagtga agtcaaactt gaagtggaaa ctgtgattaa aacaggaaga
4501 catattgtcc agaaacagca aacggacaac ccaaaaggga tggatgagca gctgacttcc
4561 ctgaaggttc tttacaatga cctgggcgca caggtgacag aaggaaaaca ggatctggaa
4621 agagcatcac agttggcccg gaaaatgaag aaagaggctg cttctctctc tgaatggctt
4681 tctgctactg aaactgaatt ggtacagaag tccacttcag aaggtctgct tggtgacttg
4741 gatacagaaa tttcctgggc taaaaatgtt ctgaaggatc tggaaaagag aaaagctgat
4801 ttaaatacca tcacagagag tagtgctgcc ctgcaaaact tgattgaggg cagtgagcct
4861 attttagaag agaggctctg cgtccttaac gctgggtgga gccgagttcg tacctggact
4921 gaagattggt gcaatacctt gatgaaccat cagaaccagc tagaaatatt tgatgggaac
4981 gtggctcaca taagtacctg gctttatcaa gctgaagctc tattggatga aattgaaaag
5041 aaaccaacaa gtaaacagga agaaattgtg aagcgtttag tatctgagct ggatgatgcc
5101 aacctccagg ttgaaaatgt ccgcgatcaa gcccttattt tgatgaatgc ccgtggaagc
5161 tcaagcaggg agcttgtaga accaaagtta gctgagctga ataggaactt tgaaaaggtg
5221 tctcaacata tcaaaagtgc caaattgcta attgctcagg aaccattata ccaatgtttg
5281 gtcaccactg aaacatttga aactggtgtg cctttctctg acttggaaaa attagaaaat
5341 gacatagaaa atatgttaaa atttgtggaa aaacacttgg aatccagtga tgaagatgaa
5401 aagatggatg aggagagtgc ccagattgag gaagttctac aaagaggaga agaaatgtta
5461 catcaaccta tggaagataa taaaaaagaa aagatccgtt tgcaattatt acttttgcat
5521 actagataca acaaaattaa ggcaatccct attcaacaga ggaaaatggg tcaacttgct
5581 tctggaatta gatcatcact tcttcctaca gattatctgg ttgaaattaa caaaatttta
5641 ctttgcatgg atgatgttga attatcgctt aatgttccag agctcaacac tgctatttac
5701 gaagacttct cttttcagga agactctctg aagaatatca aagaccaact ggacaaactt
5761 ggagagcaga ttgcagtcat tcatgaaaaa cagccagatg tcatccttga agcctctgga
5821 cctgaagcca ttcagatcag agatacactt actcagctga atgcaaaatg gggacagaatt
5881 aatagaatgt acagtgatcg gaaaggttgt tttgacaggg caatggaaga atggagacag
5941 ttccattgtg accttaatga cctcacacag tggataacag aggctgaaga attactggtt
6001 gatacctgtg ctccaggtgg cagcctggac ttagagaaag ccaggataca tcagcaggaa
6061 cttgaggtgg gcatcagcag ccaccagccc agttttgcag cactaaaccg aactggggat
6121 gggattgtgc agaaactctc ccaggcagat ggaagcttct tgaaagaaaa actggcaggt
6181 ttaaaccaac gctgggatgc aattgttgca gaagtgaagg ataggcagcc aaggctaaaa
6241 ggagaaagta agcaggtgat gaagtacagg catcagctag atgagattat ctgttggtta
6301 acaaaggctg agcatgctat gcaaaagaga tcaaccaccg aattgggaga aaacctgcaa
6361 gaattaagag acttaactca agaaatggaa gtacatgctg aaaaactcaa atggctgaat
6421 agaactgaat tggagatgct ttcagataaa agtctgagtt tacctgaaag ggataaaatt
6481 tcagaaagct taaggactgt aaatatgaca tggaataaga tttgcagaga ggtgcctacc
6541 accctgaagg aatgcatcca ggagcccagt tctgtttcac agacaaggat tgctgctcat
6601 cctaatgtcc aaaaggtggt gctagtatca tctgcgtcag atattcctgt tcagtctcat
6661 cgtacttcgg aaatttcaat tcctgctgat cttgataaaa ctataacaga actagccgac
6721 tggctggtat taatcgacca gatgctgaag tccaacattg tcactgttgg ggatgtagaa
6781 gagatcaata agaccgtttc ccgaatgaaa attacaaagg ctgacttaga acagcgccat
6841 cctcagctgg attatgtttt tacattggca cagaatttga aaaataaagc ttccagttca
6901 gatatgagaa cagcaattac agaaaaattg gaaagggtca agaaccagtg ggatggcacc
6961 cagcatggcg ttgagctaag acagcagcag cttgaggaca tgattattga cagtcttcag
7021 tgggatgacc atagggagga gactgaagaa ctgatgagaa aatatgaggc tcgactctat
7081 attcttcagc aagcccgacg ggatccactc accaaacaaa tttctgataa ccaaatactg
7141 cttcaagaac tgggtcctgg agatggtatc gtcatggcgt tcgataacgt cctgcagaaa
7201 ctcctggagg aatatgggag tgatgacaca aggaatgtga aagaaaccac agagtactta
7261 aaaacatcat ggatcaatct caaacaaagt attgctgaca gacagaacgc cttggaggct
7321 gagtggagga cggtgcaggc ctctcgcaga gatctggaaa acttcctgaa gtggatccaa
7381 gaagcagaga ccacagtgaa tgtgcttgtg gatgcctctc atcgggagaa tgctcttcag
7441 gatagtatct tggccaggga actcaaacag cagatgcagg acatccaggc agaaattgat

FIGURE 3 (cont.)

```
7501 gcccacaatg acatatttaa aagcattgac ggaaacaggc agaagatggt aaaagctttg
7561 ggaaattctg aagaggctac tatgcttcaa catcgactgg atgatatgaa ccaaagatgg
7621 aatgacttaa aagcaaaatc tgctagcatc agggcccatt tggaggccag cgctgagaag
7681 tggaacaggt tgctgatgtc cttagaagaa ctgatcaaat ggctgaatat gaaagatgaa
7741 gagcttaaga aacaaatgcc tattggagga gatgttccag ccttacagct ccagtatgac
7801 cattgtaagg ccctgagacg ggagttaaag gagaaagaat attctgtcct gaatgctgtc
7861 gaccaggccc gagttttctt ggctgatcag ccaattgagg cccctgaaga gccaagaaga
7921 aacctacaat caaaaacaga attaactcct gaggagagag cccaaaagat tgccaaagcc
7981 atgcgcaaac agtcttctga agtcaaagaa aaatgggaaa gtctaaatgc tgtaactagc
8041 aattggcaaa agcaagtgga caaggcattg gagaaactca gagacctgca gggagctatg
8101 gatgacctgg acgctgacat gaaggaggca gagtccgtgc ggaatggctg gaagcccgtg
8161 ggagacttac tcattgactc gctgcaggat cacattgaaa aaatcatggc atttagagaa
8221 gaaattgcac caatcaactt taaagttaaa acggtgaatg atttatccag tcagctgtct
8281 ccacttgacc tgcatccctc tctaaagatg tctcgccagc tagatgacct taatatgcga
8341 tggaaacttt tacaggtttc tgtggatgat cgccttaaac agcttcagga agcccacaga
8401 gattttggac catcctctca gcattttctc tctacgtcag tccagctgcc gtggcaaaga
8461 tccatttcac ataataaagt gccctattac atcaaccatc aaacacagac cacctgttgg
8521 gaccatccta aaatgaccga actctttcaa tcccttgctg acctgaataa tgtacgtttt
8581 tctgcctacc gtacagcaat caaaatccga agactacaaa aagcactatg tttggatctc
8641 ttagagttga gtacaacaaa tgaaattttc aaacagcaca agttgaacca aaatgaccag
8701 ctcctcagtg ttccagatgt catcaactgt ctgacaacaa cttatgatgg acttgagcaa
8761 atgcataagg acctggtcaa cgttccactc tgtgttgata tgtgtctcaa ttggttgctc
8821 aatgtctatg acacgggtcg aactggaaaa attagagtgc agagtctgaa gattggatta
8881 atgtctctct ccaaaggtct cttggaagaa aaatacagat atctctttaa ggaagttgcg
8941 gggccgacag aaatgtgtga ccagaggcag ctgggcctgt tacttcatga tgccatccag
9001 atcccccggc agctaggtga agtagcagct tttggaggca gtaatattga gcctagtgtt
9061 cgcagctgct tccaacagaa taacaataaa ccagaaataa gtgtgaaaga gtttatagat
9121 tggatgcatt tggaaccaca gtccatggtt tggctcccag ttttacatcg agtggcagca
9181 gcggagactg caaaacatca ggccaaatgc aacatctgta aagaatgtcc aattgtcggg
9241 ttcaggtata gaagccttaa gcattttaac tatgatgtct gccagagttg tttcttttcg
9301 ggtcgaacag caaaaggtca caaattacat tacccaatgg tggaatattg tatacctaca
9361 acatctgggg aagatgtacg agacttcaca aaggtactta agaacaagtt caggtcgaag
9421 aagtactttg ccaaacaccc tcgacttggt tacctgcctg tccagacagt tcttgaaggt
9481 gacaacttag agactcctat cacactcatc agtatgtggc cagagcacta tgacccctca
9541 caatctcctc aactgtttca tgatgacacc cattcaagaa tagaacaata tgccacacga
9601 ctggcccaga tggaaaggac taatgggtct tttctcactg atagcagctc caccacagga
9661 agtgtggaag acgagcacgc cctcatccag cagtattgcc aaacactcgg aggagagtcc
9721 ccagtgagcc agccgcagag cccagctcag atcctgaagt cagtagagag ggaagaacgt
9781 ggagaactgg agaggatcat tgctgacctg gaggaagaac aaagaaatct acaggtggag
9841 tatgagcagc tgaaggacca gcacctccga agggggctcc ctgtcggttc accgccagag
9901 tcgattatat ctccccatca cacgtctgag gattcagaac ttatagcaga agcaaaactc
9961 ctcaggcagc acaaaggtcg gctggaggct aggatgcaga ttttagaaga tcacaataaa
10021 cagctggagt ctcagctcca ccgcctccga cagctgctgg agcagcctga atctgattcc
10081 cgaatcaatg gtgtttcccc atgggcttct cctcagcatt ctgcactgag ctactcgctt
10141 gatccagatg cctccggccc acagttccac caggcagcgg gagaggacct gctggcccca
10201 ccgcacgaca ccagcacgga tctcacggag gtcatggagc agattcacag cacgtttcca
10261 tcttgctgcc caaatgttcc cagcaggcca caggcaatgt ga
```

FIGURE 4 (Mouse Utrophin cDNA, Acc. No. Y12229, SEQ ID NO:4)

```
   1 atggccaagt atggggacct tgaagccagg cctgatgatg ggcagaacga attcagtgac
  61 atcattaagt ccagatctga tgaacacaat gatgtacaga agaaaacctt taccaaatgg
 121 ataaacgctc gattttccaa gagtgggaaa ccacccatca gtgatatgtt ctcagacctc
 181 aaagatggga gaaagctctt ggatcttctc gaaggcctca caggaacatc attgccaaag
 241 gaacgtggtt ccacaagggt gcatgcctta aacaatgtca accgagtgct acaggtttta
 301 catcagaaca atgtggactt ggtgaatatt ggaggcacgg acattgtggc tggaaatccc
 361 aagctgactt tagggttact ctggagcatc attctgcact ggcaggtgaa ggatgtcatg
 421 aaagatatca tgtcagacct gcagcagaca aacagcgaga agatcctgct gagctgggtg
 481 cggcagacca ccaggcccta cagtcaagtc aacgtcctca acttcaccac cagctggacc
 541 gatggactcg cgttcaacgc cgtgctccac cggcacaaac cagatctctt cgactgggac
 601 gagatggtca aaatgtcccc aattgagaga cttgaccatg cttttgacaa ggcccacact
 661 tctttgggaa ttgaaaagct cctaagtcct gaaactgttg ctgtgcatct ccctgacaag
 721 aaatccataa ttatgtattt aacgtctctg tttgaggtgc ttcctcagca agtcacgata
 781 gatgccatcc gagaggtgga gactctccca aggaagtata agaaagaatg tgaagaggaa
 841 gaaattcata tccagagtgc agtgctggca gaggaaggcc agagtccccg agctgagacc
 901 cctagcaccg tcactgaagt ggacatggat ttggacagct accagatagc gctagaggaa
 961 gtgctgacgt ggctgctgtc cgcggaggac acgttccagg agcaacatga catttctgat
1021 gatgtcgaag aagtcaaaga gcagtttgct acccatgaaa cttttatgat ggagctgaca
1081 gcacaccaga gcagcgtggg gagcgtcctg caggctggca accagctgat gacacaaggg
1141 actctgtcca gagaggagga gtttgagatc caggaacaga tgaccttgct gaatgcaagg
1201 tgggaggcgc tccgggtgga gagcatggag aggcagtccc ggctgcacga cgctctgatg
1261 gagctgcaga agaaacagct gcagcagctc tcaagctggc tggccctcac agaagagcgc
1321 attcagaaga tggagagcct cccgctgggt gatgacctgc cctccctgca gaagctgctt
1381 caagaacata aaagtttgca aaatgacctt gaagctgaac aggtgaaggt aaattcctta
1441 actcacatgg tggtgattgt ggatgaaaac agtggggaga gtgccacagc tcttctggaa
1501 gatcagttac agaaactggg tgagcgctgg acagctgtat gccgctggac tgaagaacgt
1561 tggaacaggt tgcaagaaat cagtattctg tggcaggaat tattggaaga gcagtgtctg
1621 ttggaggctt ggctcaccga aaaggaagag gctttggata aagttcaaac cagcaacttt
1681 aaagaccaga aggaactaag tgtcagtgtc cggcgtctgg ctatattgaa ggaagacatg
1741 gaaatgaaga ggcagactct ggatcaactg agtgagattg gccaggatgt gggccaatta
1801 ctcagtaatc ccaaggcatc taagaagatg aacagtgact ctgaggagct aacacagaga
1861 tgggattctc tggttcagag actcgaagac tcttctaacc aggtgactca ggcggtagcg
1921 aagctcggca tgtcccagat tccacagaag gacctattgg agaccgttca tgtgagagaa
1981 caagggatgg tgaagaagcc caagcaggaa ctgcctcctc ctccccccacc aaagaagaga
2041 cagattcacg tggacgtgga ggccaagaaa aagtttgatg ctataagtac agagctgctg
2101 aactggattt tgaaatcaaa gactgccatt cagaacacag agatgaaaga atataagaag
2161 tcgcaggaga cctcaggaat gaaaaagaaa ttgaagggat tagagaaaga acagaaggaa
2221 aatctgcccc gactggacga actgaatcaa accggacaaa ccctccggga gcaaatggga
2281 aaagaaggcc ttccactgaa agaagtaaac gatgttctgg aaagggtttc gttggagtgg
2341 aagatgatat ctcagcagct agaagatctg ggaaggaaga tccagctgca ggaagatata
2401 aatgcttatt ttaagcagct tgatgccatt gaggagacca tcaaggagaa ggaagagtgg
2461 ctgaggggca cacccatttc tgaatcgccc cggcagccct tgccaggctt aaaggattct
2521 tgccagaggg aactgacaga tctccttggc cttcacccca gaattgagac gctgtgtgca
2581 agctgttcag ccctgaagtc tcagccctgt gtcccaggtt ttgtccagca gggttttgac
2641 gaccttcgac atcattacca ggctgttgcg aaggctttag aggaatacca acaacaacta
2701 gaaaatgagc tgaagagcca gcctggaccc gagtatttgg acacactgaa taccctgaaa
2761 aaaatgctaa gcgagtcaga aaaggcggcc caggcctctc tgaatgccct gaacgatccc
2821 atagcggtgg agcaggccct gcaggagaaa aaggcccttg atgaaaccct tgagaatcag
2881 aaacatacgt tacataagct ttcagaagaa acgaagactt tggagaaaaa tatgcttcct
2941 gatgtgggga aaatgtataa acaagaattt gatgatgtcc aaggcagatg gaataaagta
3001 aagaccaagg tttccagaga cttacacttg ctcgaggaaa tcacccccag actccgagat
3061 tttgaggctg attcagaagt cattgagaag tgggtgagtg gcatcaaaga cttcctcatg
3121 aaagaacagg ctgcccaagg agacgctgct gcgcagagcc agcttgacca atgtgctacg
3181 tttgctaatg aaatcgaaac catcgagtca tctctgaaga acatgaggga agtagagact
3241 agccttcaga ggtgtccagt cactggagtc aagacatggg tacaggcaag actagtggat
3301 taccaatccc aactggagaa attcagcaaa gagattgcta ttcaaaaaag caggctgtta
3361 gatagtcaag aaaaagccct gaacttgaaa aaggatttgg ctgagatgca ggagtggatg
3421 gcacaggctg aagaggacta cctggagagg gacttcgagt acaaatctcc agaagaactc
3481 gagagtgcgg tggaggaaat gaagagggca aaagaggatg tgctgcagaa ggaggtgagg
3541 gtgaaaattc tgaaggacag catcaagctg gtggctgcca aggtgccctc tggtggccag
3601 gagttgacgt cggaattcaa cgaggtgctg gagagctacc agcttctgtg caatagaatt
```

FIGURE 4 (cont.)

```
3661 cgagggaagt gccacacact ggaggaggtc tggtcttgct gggtggagct gcttcactat
3721 ctggacctgg agaccacgtg gttgaacacc ttggaggagc gcgtgaggag cacggaggcc
3781 ctgcctgaga gggcagaagc tgttcatgaa gctctggagt ctcttgagtc tgttttgcgc
3841 catccagcgg ataatcgcac ccagattcgg gaacttgggc agactctgat tgatggtgga
3901 atcctggatg acataatcag cgagaagctg gaggctttta acagccgcta cgaagagctg
3961 agtcacttgg cggagagcaa acagatttct ttggagaagc aactccaggt cctccgcgaa
4021 actgaccaca tgcttcaggt gctgaaggag agcctggggg agctggacaa acagcttacc
4081 acatacctga cggacaggat cgatgccttc caactgccac aggaagctca gaagatccaa
4141 gccgaaatct cagcccatga gctcaccctg gaggagctga ggaagaatgt gcgctcccag
4201 cccccgacgt cccctgaggg cagggccacc agaggaggaa gtcagatgga catgctacag
4261 aggaaacttc gagaggtctc caccaaattc cagcttttcc agaagcccgc aaatttcgag
4321 cagcggatgc tggactgcaa gcgtgtgttg gagggagtga aggccgagct tcatgtcctc
4381 gatgtgaggg atgtggaccc tgatgtcatt caggcccact tggacaagtg catgaaacta
4441 tataaaacgt tgagtgaagt caaacttgaa gttgagactg tcatcaaaac agggaggcac
4501 attgtccaga agcagcagac ggacaacccg aaaagcatgg acgaacagct tacatctctg
4561 aaagtcctct acaatgacct gggcgcacag gtgacagaag ggaagcaaga cctggaaaga
4621 gcctcacagc tgtccaggaa gatgaagaag gaggctgccg tcctctctga atggctctct
4681 gccacagagg cagaactagt gcagaaatcc acatcagaag gcgtgattgg tgacctggac
4741 acagaaatct cctgggctaa aagtattctc aaggatctgg aaaagaggaa agttgactta
4801 aatggcatta cagagagcag tgctgccctt cagcacttgg tcttgggcag tgagtctgtt
4861 ctggaagaga acctctgtgt gctcaatgct ggatggagcc gagtgcggac gtggaccgaa
4921 gactggtgca acaccttgct gaaccatcaa aaccagctgg agctatttga tggacacgtc
4981 gctcacatca gtacctggct ctatcaagca gaagctctgc tggatgagat cgaaaagaaa
5041 ccagcgagta aacaggaaga aattgtgaag cgtttactgt ctgaattgga tgatgccagc
5101 ctccaggttg agaatgttcg ggaacaagcc atcatcttgg tgaatgctcg tggaagcgcc
5161 agcagggaac tcgtggaacc aaaattagcc gagctgagca ggaactttga aaaggtgtcc
5221 cagcacataa agagcgcccg aatgctgatt ggtcaggacc cttcatccta ccaaggcttg
5281 gaccctgctg gaactgttca agctgctgag tctttctctg acttggaaaa cttagaacaa
5341 gacatagaaa acatgttgaa agttgtggaa aagcacttgg accccaataa cgatgagaag
5401 atggatgagg agcaagccca gattgaggaa gttctacaaa gaggggagca tttgttacat
5461 gaacctatgg aggacagtaa gaaagaaaag atccgcttgc agttgttact tttgcatact
5521 cgttacaaca aaattaagac aatccctatc cagcagagaa aaacaattcc agtttcttct
5581 ggaattacat catcagccct ccctgcagat tatttggttg aaattaataa aatttactc
5641 actctggatg acattgaatt atcacttaat atgccggagc taaacaccac tgtctacaaa
5701 gacttctctt tccaggaaga ctctctgaag agtatcaaag gtcaactgga cagacttgga
5761 gagcagattg cagttgttca cgagaagcag ccggatgtca tcgtggaagc ctctggccct
5821 gaggccattc agatcaggga catgctcgct cagctgaacg caaaatggga ccgagtgaat
5881 agagtgtaca gtgatcggag agggtccttt gccagggctg tggaggaatg gaggcagttc
5941 caccatgacc ttgatgacct tacacagtgg ctatctgaag ctgaagacct gctggtagac
6001 acttgtgctc cagatggtag cctggacctg gagaaagcca gggcacagca gctggaactg
6061 gaagagggcc tcagcagcca ccagcccagc ctgatcaagg ttaaccgaaa gggggaggac
6121 cttgttcaga gactccgccc ctcggaggca agcttcctga aggagaagct ggcaggtttc
6181 aaccagcgct ggagcactct tgtagctgag gtggaggctt tgcagcccag gctaaaagga
6241 gaaagtcagc aggtgttggg gtataagaga cggctagatg aggtcacctg ctggttaacg
6301 aaagtggaga gtgctgtgca gaagagatca acccctgacc cggaagaaag cccacaggaa
6361 ttaacagatt tagcccaaga gacggaagtt caagctgaaa acattaagtg gctgaacaga
6421 gcagaactgg aaatgctttc agacaaaaat ctgagtttgc gtgaaagaga gaaactttcg
6481 gaaagtttaa agaatgtaaa cacaacatgg accaaggtat gcagagaagt gcctagcctc
6541 ctgaagacac gcacccaaga cccctgctct gccccacaga tgaggatggc tgctcatccc
6601 aacgtccaaa aggtggtgct agtatcatct gcatcagatg ctcctctgcg tggcggcctg
6661 gaaatctcgg ttcctgctga tttggataaa accatcacag aactggctga ctggctggta
6721 ttgatcgacc aaatgctgaa gtccaacatt gtcactgtgg gggacgtgaa agagatcaat
6781 aagacagttt cccggatgaa aatcacaaag gctgatttag aacaacgcca tcctcagctt
6841 gattgtgtat ttacgttggc ccaaaatttg aaaaacaaag cttccagttc agatgtgaga
6901 acagcaatca cagaaaaatt ggaaaagctg aagacccagt gggagagtac tcagcatggt
6961 gtggagctgc ggcggcagca gctggaggac atggttgtgg acagcctgca gtgggacgac
7021 cacagggaag agactgaaga gctcatgaga aaatacgagg ctcgcttcta catgctgcag
7081 caggcccgcc gggacccact tagcaaacaa gtttctgata atcaactatt gcttcaagag
7141 ctggggtctg gcgatggtgt catcatggcg tttgataatg tcctgcagaa acttctggaa
7201 gaatacagtg gcgatgacac aaggaatgtg gaagaaacca cggagtactt gaaaacatca
7261 tgggtcaatc tcaaacaaag catcgctgat agacagagtg ccttggaggc tgagctacag
7321 acagtgcaga cttctcgtag agacctggag aactttgtca agtggcttca ggaagcagaa
7381 accacagcaa atgtgctggc cgatgcctct cagcgggaga atgctcttca ggacagtgtc
7441 ctggcccggc agctccgaca gcagatgctg gacatccagg cagaaattga tgcccacaat
```

FIGURE 4 (cont.)

```
7501 gacatattta aaagcatcga tggaaaccgg cagaagatgg tgaaagctct ggggaattct
7561 gaggaagcaa caatgcttca acatcgactg gatgacatga accaaagatg gaatgatttg
7621 aaggcaaaat ctgctagcat cagggcccat ttggaggcca gtgctgagaa atggaaccgg
7681 ttgctggcat cgctggaaga gctgatcaaa tggctcaata tgaaagatga ggagcttaag
7741 aagcagatgc ccattggagg ggacgtccct gccttacagc tccagtatga ccactgcaag
7801 gtgctgagac gtgagctaaa ggagaaagag tattctgtgc tgaacgccgt agatcaagct
7861 cgagtttttc tggctgatca gccaatagag gcccccgaag aaccaagaag aaacccacaa
7921 tcaaagacag agttgactcc tgaggagaga gcccagaaga tcgccaaagc catgcgcaag
7981 cagtcttctg aagtccgaga gaagtgggaa aatctaaatg ctgtcactag caactggcaa
8041 aagcaagtag ggaaggcgtt agagaaactc cgagacctgc agggagctat ggacgacctg
8101 gacgcagaca tgaaggaggt ggaggctgtg cggaatggct ggaagcccgt gggagacctg
8161 cttatagact ccctgcagga tcacatcgag aaaaccctgg cgtttagaga agaaattgca
8221 ccaatcaact taaaagtaaa aacaatgaat gacctgtcca gtcagctgtc tccacttgac
8281 ttgcatccat ctctaaagat gtctcgccag ctggatgacc ttaatatgcg atggaaactt
8341 ctacaggttt ccgtggacga tcgccttaag cagctccagg aagcccacag agattttggg
8401 ccatcttctc aacactttct gtccacttca gtccagctgc cgtggcagag atccatttca
8461 cataataaag tgccctatta catcaaccat caaacacaga caacctgttg ggatcatcct
8521 aaaatgactg agctcttcca atcccttgct gatctgaata atgtacgttt ctctgcctac
8581 cgcacagcaa tcaaaattcg aaggctgcaa aaagcattat gtctggatct cttagagctg
8641 aatacgacga atgaagtttt caagcagcac aaactgaacc aaaatgatca gctcctgagt
8701 gtcccagacg tcatcaactg tctgaccacc acttacgatg ggcttgagca gctgcacaag
8761 gacttggtca atgttccact ctgcgtcgat atgtgtctca actggctgct caacgtatac
8821 gacacgggcc ggactggaaa aattcgggta cagagtctga agattggatt gatgtctctc
8881 tccaaaggcc tcttagaaga gaaatacaga tgtctcttta aggaggtggc agggccaact
8941 gagatgtgtg accagcggca gcttggcctg ctacttcacg atgccatcca gatccctagg
9001 cagctggggg aagtagcagc ctttgggggc agtaacattg agcccagtgt ccgcagctgc
9061 ttccagcaga ataacaacaa gccagaaatc agtgtgaagg agtttataga ctggatgcat
9121 ttggaacccc agtccatggt gtggttgccg gttctgcatc gggtcgcagc tgctgagact
9181 gcaaaacatc aggccaaatg caacatctgc aaagaatgcc cgattgttgg gttcagatac
9241 aggagcctaa agcattttaa ttatgatgtc tgccagagtt gcttcttttc tggaagaaca
9301 gcaaagggcc acaagttaca ttacccgatg gtagaatact gcataccgac aacatctggg
9361 gaagatgtga gagatttcac taaggtgctg aagaacaagt tcaggtccaa gaaatatttt
9421 gccaaacatc ctcggcttgg ctacctgcct gtccagaccg tgctggaagg ggacaactta
9481 gaaactccta tcacgctcat cagtatgtgg ccagagcact atgacccctc ccagtccct
9541 cagctgtttc atgatgacac ccactcaaga atagagcaat acgctacacg actggcccag
9601 atggaaagga caaacgggtc cttcctaact gatagcagct ctacaacagg aagcgtggag
9661 gatgagcatg ccctcatcca gcagtactgc cagaccctgg gcggggagtc acctgtgagt
9721 cagccgcaga gtccagctca gatcctgaag tccgtggaga gggaagagcg tggggaactg
9781 gagcggatca ttgctgactt ggaggaagag caaagaaatc tgcaggtgga gtatgagcag
9841 ctgaaggagc agcacctaag aaggggtctc cctgtgggct cccctccaga ctccatcgta
9901 tctcctcacc acacatctga ggactcagaa cttatagcag aagctaaact cctgcggcag
9961 cacaaagggc ggctggaggc gaggatgcaa attttggaag atcacaataa acagctggag
10021 tctcagctgc accgcctcag acagctcctg gagcagcctg actctgactc ccgcatcaat
10081 ggtgtctccc cctgggcttc cccacagcat tctgcattga gctactcact tgacactgac
10141 ccaggcccac agttccacca ggcagcatct gaggacctgc tggccccacc tcacgacact
10201 agcacggacc tcacggacgt gatggagcag atcaacagca cgtttccctc ttgcagctca
10261 aatgtcccca gcaggccaca ggcaatgtga gcatctatcc agccagccaa catttcccga
10321 ccttcagtat tgccctcttc tgcaaatgcc aatcccaaga cccattcaac cccaaagctc
10381 cgtggctcca cgacacaagc tgttgagtgc ttactgggtg ttctactgag ggaaccaaac
10441 actgactatc caaagatatt ttggttttct aataacgtat attattgttt tctttctccc
10501 ctttctatgc aactgtaaat taatgaacag agaagtattt ggaggtggta aagcatttgt
10561 cactgatttg tataatatat acagccatgg gaaagtgggt gggggctttc taatatgaaa
10621 ctgtcttttt aataaccaag agaaaaaatt gcataagaat tagaccactt tacattatta
10681 cattccttct gctgttcaca ttaaccttgt acaataactt cacttattat ttgactgttt
10741 taccattatg ttttggttat ttataaattt atcagccata ccaaacgaat agattctatg
10801 tatttggttt ctataatctg gccaaattcc taagttcata tatttgaatc aaatatttta
10861 catatgtgga gtaggcaggc attctgaaga tactatttaa ctttagttga cgtcacacac
10921 accatccttt agtaaccact ggatgactac actaaaaatc ctgtggactt taacggcaag
10981 ctgctggggt atttttcctc ctgtttttat tccttttttg taagtagatc ttgacgtctt
11041 tatttatttc atcttgcaat ctctataata aagaagactg tattgtaata gtcccc
```

6023848.100600

FIGURE 5

SEQ ID NO:5 (5' UTR, 1-208))
1 gggattccct cactttcccc ctacaggact cagatctggg aggcaattac cttcggagaa
61 aaacgaatag gaaaaactga agtgttactt tttttaaagc tgctgaagtt tgttggtttc
121 tcattgtttt taagcctact ggagcaataa agtttgaaga acttttacca ggttttttt
181 atcgctgcct tgatatacac ttttcaaa

SEQ ID NO:6 (N terminus, 209-964)
209 at gctttggtgg gaagaagtag aggactgtta
241 tgaaagagaa gatgttcaaa agaaaacatt cacaaaatgg gtaaatgcac aatttttctaa
301 gtttgggaag cagcatattg agaacctctt cagtgaccta caggatggga ggcgcctcct
361 agacctcctc gaaggcctga cagggcaaaa actgccaaaa gaaaaaggat ccacaagagt
421 tcatgccctg aacaatgtca acaaggcact gcgggttttg cagaacaata atgttgattt
481 agtgaatatt ggaagtactg acatcgtaga tggaaatcat aaactgactc ttggtttgat
541 ttggaatata atcctccact ggcaggtcaa aaatgtaatg aaaaatatca tggctggatt
601 gcaacaaacc aacagtgaaa agattctcct gagctgggtc cgacaatcaa ctcgtaatta
661 tccacaggtt aatgtaatca acttcaccac cagctggtct gatggcctgg ctttgaatgc
721 tctcatccat agtcataggc cagacctatt tgactggaat agtgtggttt gccagcagtc
781 agccacacaa cgactggaac atgcattcaa catcgccaga tatcaattag gcatagagaa
841 actactcgat cctgaagatg ttgataccac ctatccagat aagaagtcca tcttaatgta
901 catcacatca ctcttccaag ttttgcctca acaagtgagc attgaagcca tccaggaagt
961 ggaa

SEQ ID NO:7 (Hinge 1, 965-1219)
965 atgttg ccaaggccac ctaaagtgac taaagaagaa cattttcagt tacatcatca
1021 aatgcactat tctcaacaga tcacggtcag tctagcacag ggatatgaga gaacttcttc
1081 ccctaagcct cgattcaaga gctatgccta cacacaggct gcttatgtca ccacctctga
1141 ccctacacgg agcccatttc cttcacagca tttggaagct cctgaagaca agtcatttgg
1201 cagttcattg atggagagt

SEQ ID NO:8 (Repeat 1, 1220-1546)
1220 g aagtaaacct ggaccgttat caaacagctt tagaagaagt
1261 attatcgtgg cttctttctg ctgaggacac attgcaagca caaggagaga tttctaatga
1321 tgtggaagtg gtgaaagacc agtttcatac tcatgagggg tacatgatgg atttgacagc
1381 ccatcagggc cgggttggta atattctaca attgggaagt aagctgattg gaacaggaaa
1441 attatcagaa gatgaagaaa ctgaagtaca agagcagatg aatctcctaa attcaagatg
1501 ggaatgcctc agggtagcta gcatggaaaa acaaagcaat ttacat

SEQ ID NO:9 (Repeat 2, 1547-1879)
1547 agag ttttaatgga
1561 tctccagaat cagaaactga aagagttgaa tgactggcta acaaaaacag aagaaagaac
1621 aaggaaaatg gaggaagagc ctcttggacc tgatcttgaa gacctaaaac gccaagtaca
1681 acaacataag gtgcttcaag aagatctaga acaagaacaa gtcagggtca attctctcac
1741 tcacatggtg gtggtagttg atgaatctag tggagatcac gcaactgctg ctttggaaga
1801 acaacttaag gtattgggag atcgatgggc aaacatctgt agatggacag aagaccgctg
1861 ggttctttta caagacatc

SEQ ID NO:10 (Repeat 3, 1880-2212)
1880 c ttctcaaatg gcaacgtctt actgaagaac agtgcctttt
1921 tagtgcatgg ctttcagaaa aagaagatgc agtgaacaag attcacacaa ctggctttaa
1981 agatcaaaat gaaatgttat caagtcttca aaaactggcc gttttaaaag cggatctaga
2041 aaagaaaaag caatccatgg gcaaactgta ttcactcaaa caagatcttc tttcaacact
2101 gaagaataag tcagtgaccc agaagacgga agcatggctg gataactttg cccggtgttg
2161 ggataattta gtccaaaaac ttgaaaagag tacagcacag atttcacagg ct

SEQ ID NO:11 (Hinge 2, 2213-2359)
2213 gtcaccac
2221 cactcagcca tcactaacac agacaactgt aatggaaaca gtaactacgg tgaccacaag
2281 ggaacagatc ctggtaaagc atgctcaaga ggaacttcca ccaccacctc cccaaaagaa
2341 gaggcagatt actgtggat

FIGURE 6

SEQ ID NO:12 (Repeat 4, 2360-2692)

```
2360 t ctgaaattag gaaaaggttg gatgttgata taactgaact
2401 tcacagctgg attactcgct cagaagctgt gttgcagagt cctgaatttg caatctttcg
2461 gaaggaaggc aacttctcag acttaaaaga aaaagtcaat gccatagagc gagaaaaagc
2521 tgagaagttc agaaaactgc aagatgccag cagatcagct caggccctgg tggaacagat
2581 ggtgaatgag ggtgttaatg cagatagcat caaacaagcc tcagaacaac tgaacagccg
2641 gtggatcgaa ttctgccagt tgctaagtga gagacttaac tggctggagt at
```

SEQ ID NO:13 (Repeat 5, 2693-3019)

```
2693 cagaacaa
2701 catcatcgct ttctataatc agctacaaca attggagcag atgacaacta ctgctgaaaa
2761 ctggttgaaa atccaaccca ccaccccatc agagccaaca gcaattaaaa gtcagttaaa
2821 aatttgtaag gatgaagtca accggctatc aggtcttcaa cctcaaattg aacgattaaa
2881 aattcaaagc atagccctga aagagaaagg acaaggaccc atgttcctgg atgcagactt
2941 tgtggccttt acaaatcatt ttaagcaagt cttttctgat gtgcaggcca gagagaaaga
3001 gctacagaca atttttgac
```

SEQ ID NO:14 (Repeat 6, 3020-3346)

```
3020 a ctttgccacc aatgcgctat caggagacca tgagtgccat
3061 caggacatgg gtccagcagt cagaaaccaa actctccata cctcaactta gtgtcaccga
3121 ctatgaaatc atggagcaga gactcgggga attgcaggct ttacaaagtt ctctgcaaga
3181 gcaacaaagt ggcctatact atctcagcac cactgtgaaa gagatgtcga agaaagcgcc
3241 ctctgaaatt agccggaaat atcaatcaga atttgaagaa attgagggac gctggaagaa
3301 gctctcctcc cagctggttg agcattgtca aaagctagag gagcaa
```

SEQ ID NO:15 (Repeat 7, 3347-3673)

```
3347 atga ataaactccg
3361 aaaaattcag aatcacatac aaaccctgaa gaaatggatg gctgaagttg atgtttttct
3421 gaaggaggaa tggcctgccc ttggggattc agaaattcta aaaaagcagc tgaaacagtg
3481 cagactttta gtcagtgata ttcagacaat tcagcccagt ctaaacagtg tcaatgaagg
3541 tgggcagaag ataaagaatg aagcagagcc agagtttgct tcgagacttg agacagaact
3601 caaagaactt aacactcagt gggatcacat gtgccaacag gtctatgcca gaaaggaggc
3661 cttgaaggga ggt
```

SEQ ID NO:16 (Repeat 8, 3674-4000)

```
3674 ttggaga aaactgtaag cctccagaaa gatctatcag agatgcacga
3721 atggatgaca caagctgaag aagagtatct tgagagagat tttgaatata aaactccaga
3781 tgaattacag aaagcagttg aagagatgaa gagagctaaa gaagaggccc aacaaaaaga
3841 agcgaaagtg aaactcctta ctgagtctgt aaatagtgtc atagctcaag ctccacctgt
3901 agcacaagag gccttaaaaa aggaacttga aactctaacc accaactacc agtggctctg
3961 cactaggctg aatgggaaat gcaagacttt ggaagaagtt
```

SEQ ID NO:17 (Repeat 9, 4001-4312)

```
4001 tgggcatgtt ggcatgagtt
4021 attgtcatac ttggagaaag caaacaagtg gctaaatgaa gtagaattta aacttaaaac
4081 cactgaaaac attcctggcg gagctgagga aatctctgag gtgctagatt cacttgaaaa
4141 tttgatgcga cattcagagg ataacccaaa tcagattcgc atattggcac agaccctaac
4201 agatggcgga gtcatggatg agctaatcaa tgaggaactt gagacattta attctcgttg
4261 gagggaacta catgaagagg ctgtaaggag gcaaaagttg cttgaacaga gc
```

SEQ ID NO:18 (Repeat 10, 4313-4588)

```
4313 atccagtc
4321 tgcccaggag actgaaaaat ccttacactt aatccaggag tccctcacat tcattgacaa
4381 gcagttggca gcttatattg cagacaaggt ggacgcagct caaatgcctc aggaagccca
4441 gaaaatccaa tctgatttga caagtcatga gatcagttta gaagaaatga agaaacataa
4501 tcagggaag gaggctgccc aaagagtcct gtctcagatt gatgttgcac agaaaaaatt
4561 acaagatgtc tccatgaagt ttcgatta
```

FIGURE 7

SEQ ID NO:19 (Repeat 11, 4589-4915)
4589 tt ccagaaacca gccaattttg agctgcgtct
4621 acaagaaagt aagatgattt tagatgaagt gaagatgcac ttgcctgcat tggaaacaaa
4681 gagtgtggaa caggaagtag tacagtcaca gctaaatcat tgtgtgaact tgtataaaag
4741 tctgagtgaa gtgaagtctg aagtggaaat ggtgataaag actggacgtc agattgtaca
4801 gaaaaagcag acggaaaatc ccaaagaact tgatgaaaga gtaacagctt tgaaattgca
4861 ttataatgag ctgggagcaa aggtaacaga aagaaagcaa cagttggaga aatgc

SEQ ID NO:20 (Repeat 12, 4916-5239)
4916 ttgaa
4921 attgtcccgt aagatgcgaa aggaaatgaa tgtcttgaca gaatggctgg cagctacaga
4981 tatggaattg acaaagagat cagcagttga aggaatgcct agtaatttgg attctgaagt
5041 tgcctgggga aaggctactc aaaaagagat tgagaaacag aaggtgcacc tgaagagtat
5101 cacagaggta ggagaggcct tgaaaacagt tttgggcaag aaggagacgt tggtggaaga
5161 taaactcagt cttctgaata gtaactggat agctgtcacc tcccgagcag aagagtggtt
5221 aaatcttttg ttggaatac

SEQ ID NO:21 (Repeat 13, 5240-5551)
5240 c agaaacacat ggaaactttt gaccagaatg tggaccacat
5281 cacaaagtgg atcattcagg ctgacacact tttggatgaa tcagagaaaa agaaaccсса
5341 gcaaaaagaa gacgtgctta agcgtttaaa ggcagaactg aatgacatac gcccaaaggt
5401 ggactctaca cgtgaccaag cagcaaactt gatggcaaac cgcggtgacc actgcaggaa
5461 attagtagag ccccaaatct cagagctcaa ccatcgattt gcagccattt cacacagaat
5521 taagactgga aaggcctcca ttcctttgaa g

SEQ ID NO:22 (Repeat 14, 5552-5833)
5552 gaattggag cagtttaact cagatataca
5581 aaaattgctt gaaccactgg aggctgaaat tcagcagggg gtgaatctga aagaggaaga
5641 cttcaataaa gatatgaatg aagacaatga gggtactgta aaagaattgt tgcaaagagg
5701 agacaactta caacaaagaa tcacagatga gagaaagaga gaggaaataa agataaaaca
5761 gcagctgtta cagacaaaac ataatgctct caaggatttg aggtctcaaa gaagaaaaaa
5821 ggctctagaa att

SEQ ID NO:23 (Repeat 15, 5834-6127)
5834 tctcatc agtggtatca gtacaagagg caggctgatg atctcctgaa
5881 atgcttggat gacattgaaa aaaaattagc cagcctacct gagcccagag atgaaaggaa
5941 aataaaggaa attgatcggg aattgcagaa gaagaaagag gagctgaatg cagtgcgtag
6001 gcaagctgag ggcttgtctg aggatggggc cgcaatggca gtggagccaa ctcagatcca
6061 gctcagcaag cgctggcggg aaattgagag caaatttgct cagtttcgaa gactcaactt
6121 tgcacaa

SEQ ID NO:24 (Repeat 16, 6188-6514)
6128 tct tatgtgcctt ctacttattt gactgaaatc actcatgtct cacaagccct
6241 attagaagtg gaacaacttc tcaatgctcc tgacctctgt gctaaggact ttgaagatct
6301 ctttaagcaa gaggagtctc tgaagaatat aaaagatagt ctacaacaaa gctcaggtcg
6361 gattgacatt attcatagca agaagacagc agcattgcaa agtgcaacgc ctgtggaaag
6421 ggtgaagcta caggaagctc tctcccagct tgatttccaa tgggaaaaag ttaacaaaat
6481 gtacaaggac cgacaagggc gatttgacag atct

SEQ ID NO:25 (Repeat 17, 6515-6835)
6515 gttgag aaatggcggc gttttcatta
6541 tgatataaag atatttaatc agtggctaac agaagctgaa cagtttctca gaaagacaca
6601 aattcctgag aattgggaac atgctaaata caaatggtat cttaaggaac tccaggatgg
6661 cattgggcag cggcaaactg ttgtcagaac attgaatgca actggggaag aaataattca
6721 gcaatcctca aaaacagatg ccagtattct acaggaaaaa ttgggaagcc tgaatctgcg
6781 gtggcaggag gtctgcaaac agctgtcaga cagaaaaaag aggctagaag aacaa

FIGURE 8

SEQ ID NO:26 (Repeat 18, 6836-7186)

6836 aagaa
6841 tatcttgtca gaatttcaaa gagatttaaa tgaatttgtt ttatggttgg aggaagcaga
6901 taacattgct agtatcccac ttgaacctgg aaaagagcag caactaaaag aaaagcttga
6961 gcaagtcaag ttactggtgg aagagttgcc cctgcgccag ggaattctca aacaattaaa
7021 tgaaactgga ggacccgtgc ttgtaagtgc tcccataagc ccagaagagc aagataaact
7081 tgaaaataag ctcaagcaga caaatctcca gtggataaag gtttccagag ctttacctga
7141 gaaacaagga gaaattgaag ctcaaataaa agaccttggg cagctt

SEQ ID NO:27 (Repeat 19, 7187-7489)

7187 gaaa aaaagcttga
7201 agaccttgaa gagcagttaa atcatctgct gctgtggtta tctcctatta ggaatcagtt
7261 ggaaatttat aaccaaccaa accaagaagg accatttgac gttcaggaaa ctgaaatagc
7321 agttcaagct aaacaaccgg atgtggaaga gattttgtct aaagggcagc atttgtacaa
7381 ggaaaaacca gccactcagc cagtgaagag gaagttagaa gatctgagct ctgagtggaa
7441 ggcggtaaac cgtttacttc aagagctgag ggcaaagcag cctgaccta

SEQ ID NO:28 (Hinge 3, 7490-7612)

7490 g ctcctggact
7501 gaccactatt ggagcctctc ctactcagac tgttactctg gtgacacaac ctgtggttac
7561 taaggaaact gccatctcca aactagaaat gccatcttcc ttgatgttgg ag

SEQ ID NO:29 (Repeat 20, 7613-7942)

7613 gtacctgc
7621 tctggcagat ttcaaccggg cttggacaga acttaccgac tggctttctc tgcttgatca
7681 agttataaaa tcacagaggg tgatggtggg tgaccttgag gatatcaacg agatgatcat
7741 caagcagaag gcaacaatgc aggatttgga acagaggcgt ccccagttgg aagaactcat
7801 taccgctgcc caaaatttga aaaacaagac cagcaatcaa gaggctagaa caatcattac
7861 ggatcgaatt gaaagaattc agaatcagtg ggatgaagta caagaacacc ttcagaaccg
7921 gaggcaacag ttgaatgaaa tg

SEQ ID NO:30 (Repeat 21, 7943-8269)

7943 ttaaagga ttcaacacaa tggctggaag ctaaggaaga
7981 agctgagcag gtcttaggac aggccagagc caagcttgag tcatggaagg agggtcccta
8041 tacagtagat gcaatccaaa agaaaatcac agaaaccaag cagttggcca aagacctccg
8101 ccagtggcag acaaatgtag atgtggcaaa tgacttggcc ctgaaacttc tccgggatta
8161 ttctgcagat gataccagaa aagtccacat gataacagag aatatcaatg cctcttggag
8221 aagcattcat aaaagggtga gtgagcgaga ggctgctttg gaagaaact

SEQ ID NO:31 (Repeat 22, 8270-8617)

8270 c atagattact
8281 gcaacagttc cccctggacc tggaaaagtt tcttgcctgg cttacagaag ctgaaacaac
8341 tgccaatgtc ctacaggatg ctacccgtaa ggaaaggctc ctagaagact ccaagggagt
8401 aaaagagctg atgaaacaat ggcaagacct ccaaggtgaa attgaagctc acacagatgt
8461 ttatcacaac ctggatgaaa acagccaaaa aatcctgaga tccctggaag gttccgatga
8521 tgcagtcctg ttacaaagac gtttggataa catgaacttc aagtggagtg aacttcggaa
8581 aaagtctctc aacattaggt cccatttgga agccagt

SEQ ID NO:32 (Repeat 23, 8618-9004)

8618 tct gaccagtgga agcgtctgca
8641 cctttctctg caggaacttc tggtgtggct acagctgaaa gatgatgaat taagccggca
8701 ggcacctatt ggaggcgact ttccagcagt tcagaagcag aacgatgtac atagggcctt
8761 caagagggaa ttgaaaacta aagaacctgt aatcatgagt actcttgaga ctgtacgaat
8821 atttctgaca gagcagcctt tggaaggact agagaaactc taccaggagc ccagagagct
8881 gcctcctgag gagagagccc agaatgtcac tcggcttcta cgaaagcagg ctgaggaggt
8941 caatactgag tgggaaaaat tgaacctgca ctccgctgac tggcagagaa aaatagatga
9001 gacc

FIGURE 9

SEQ ID NO:33 (Repeat 24, 9005-9328)

9005 cttgaa agactccagg aacttcaaga ggccacggat gagctggacc tcaagctgcg
9061 ccaagctgag gtgatcaagg gatcctggca gcccgtgggc gatctcctca ttgactctct
9121 ccaagatcac ctcgagaaag tcaaggcact tcgaggagaa attgcgcctc tgaaagagaa
9181 cgtgagccac gtcaatgacc ttgctcgcca gcttaccact ttgggcattc agctctcacc
9241 gtataacctc agcactctgg aagacctgaa caccagatgg aagcttctgc aggtggccgt
9301 cgaggaccga gtcaggcagc tgcatgaa

SEQ ID NO:34 (Hinge 4, 9329-9544)

9329 c ccacagggac tttggtccag catctcagca
9361 ctttctttcc acgtctgtcc agggtccctg ggagagagcc atctcgccaa acaaagtgcc
9421 ctactatatc aaccacgaga ctcaaacaac ttgctgggac catcccaaaa tgacagagct
9481 ctaccagtct ttagctgacc tgaataatgt cagattctca gcttatagga ctgccatgaa
9541 actc

SEQ ID NO:35 (Start of C terminus, 9545-10431)

9545 cgaaga ctgcagaagg ccctttgctt ggatctcttg agcctgtcag ctgcatgtga
9601 tgccttggac cagcacaacc tcaagcaaaa tgaccagccc atggatatcc tgcagattat
9661 taattgtttg accactattt atgaccgcct ggagcaagag cacaacaatt tggtcaacgt
9721 ccctctctgc gtggatatgt gtctgaactg gctgctgaat gtttatgata cgggacgaac
9781 agggaggatc cgtgtcctgt cttttaaaac tggcatcatt tccctgtgta aagcacattt
9841 ggaagacaag tacagatacc ttttcaagca agtggcaagt tcaacaggat tttgtgacca
9901 gcgcaggctg ggcctccttc tgcatgattc tatccaaatt ccaagacagt tgggtgaagt
9961 tgcatccttt gggggcagta acattgagcc aagtgtccgg agctgcttcc aatttgctaa
10021 taataagcca gagatcgaag cggccctctt cctagactgg atgagactgg aaccccagtc
10081 catggtgtgg ctgcccgtcc tgcacagagt ggctgctgca gaaactgcca agcatcaggc
10141 caaatgtaac atctgcaaag agtgtccaat cattggattc aggtacagga gtctaaagca
10201 ctttaattat gacatctgcc aaagctgctt tttttctggt cgagttgcaa aaggccataa
10261 aatgcactat cccatggtgg aatattgcac tccgactaca tcaggagaag atgttcgaga
10321 ctttgccaag gtactaaaaa acaaatttcg aaccaaaagg tattttgcga agcatccccg
10381 aatgggctac ctgccagtgc agactgtctt agagggggac aacatggaaa c

SEQ ID NO:36 (alternatively spliced exons 71-78, 10432-11254)

10432 tcccgttac
10441 tctgatcaac ttctggccag tagattctgc gcctgcctcg tcccctcagc tttcacacga
10501 tgatactcat tcacgcattg aacattatgc tagcaggcta gcagaaatgg aaaacagcaa
10561 tggatcttat ctaaatgata gcatctctcc taatgagagc atagatgatg aacatttgtt
10621 aatccagcat tactgccaaa gtttgaacca ggactccccc ctgagccagc ctcgtagtcc
10681 tgcccagatc ttgatttcct tagagagtga ggaaagaggg gagctagaga gaatcctagc
10741 agatcttgag gaagaaaaca ggaatctgca agcagaatat gaccgtctaa agcagcagca
10801 cgaacataaa ggcctgtccc cactgccgtc ccctcctgaa atgatgccca cctctcccca
10861 gagtccccgg gatgctgagc tcattgctga ggccaagcta ctgcgtcaac acaaaggccg
10921 cctggaagcc aggatgcaaa tcctggaaga ccacaataaa cagctggagt cacagttaca
10981 caggctaagg cagctgctgg agcaacccca ggcagaggcc aaagtgaatg gcacaacggt
11041 gtcctctcct tctacctctc tacagaggtc cgacagcagt cagcctatgc tgctccgagt
11101 ggttggcagt caaacttcgg actccatggg tgaggaagat cttctcagtc ctccccagga
11161 cacaagcaca gggttagagg aggtgatgga gcaactcaac aactccttcc ctagttcaag
11221 aggaagaaat acccctggaa agccaatgag agag

SEQ ID NO:37 (End of coding region, 11255-11266)

11255 gacaca atgtag

FIGURE 10

SEQ ID NO:38 (3' UTR, 11267-13957)

```
11267 gaag tcttttccac
11281 atggcagatg atttgggcag agcgatggag tccttagtat cagtcatgac agatgaagaa
11341 ggagcagaat aaatgtttta caactcctga ttcccgcatg gtttttataa tattcataca
11401 acaaagagga ttagacagta agagtttaca agaaataaat ctatattttt gtgaagggta
11461 gtggtattat actgtagatt tcagtagttt ctaagtctgt tattgttttg ttaacaatgg
11521 caggttttac acgtctatgc aattgtacaa aaaagttata agaaaactac atgtaaaatc
11581 ttgatagcta aataacttgc catttcttta tatggaacgc attttgggtt gtttaaaaat
11641 ttataacagt tataaagaaa gattgtaaac taaagtgtgc tttataaaaa aaagttgttt
11701 ataaaaaccc ctaaaaacaa aacaaacaca cacacacaca catacacaca cacacacaaa
11761 actttgaggc agcgcattgt tttgcatcct tttggcgtga tatccatatg aaattcatgg
11821 ctttttcttt ttttgcatat taaagataag acttcctcta ccaccacacc aaatgactac
11881 tacacactgc tcatttgaga actgtcagct gagtggggca ggcttgagtt ttcatttcat
11941 atatctatat gtctataagt atataaatac tatagttata tagataaaga gatacgaatt
12001 tctatagact gacttttcc atttttaaa tgttcatgtc acatcctaat agaaagaaat
12061 tacttctagt cagtcatcca ggcttacctg cttggtctag aatggatttt tcccggagcc
12121 ggaagccagg aggaaactac accacactaa aacattgtct acagctccag atgtttctca
12181 ttttaaacaa ctttccactg acaacgaaag taaagtaaag tattggattt ttttaaaggg
12241 aacatgtgaa tgaatacaca ggacttatta tatcagagtg agtaatcggt tggttggttg
12301 attgattgat tgattgatac attcagcttc ctgctgctag caatgccacg atttagattt
12361 aatgatgctt cagtggaaat caatcagaag gtattctgac cttgtgaaca tcagaaggta
12421 ttttttaact cccaagcagt agcaggacga tgatagggct ggagggctat ggattcccag
12481 cccatccctg tgaaggagta ggccactctt taagtgaagg attggatgat tgttcataat
12541 acataaagtt ctctgtaatt acaactaaat tattatgccc tcttctcaca gtcaaaagga
12601 actgggtggt ttggtttttg ttgctttttt agatttattg tcccatgtgg gatgagtttt
12661 taaatgccac aagacataat ttaaaataaa taaactttgg gaaaaggtgt aagacagtag
12721 ccccatcaca tttgtgatac tgacaggtat caacccagaa gcccatgaac tgtgtttcca
12781 tcctttgcat ttctctgcga gtagttccac acaggtttgt aagtaagtaa gaaagaaggc
12841 aaattgattc aaatgttaca aaaaaaccct tcttggtgga ttagacaggt taaatatata
12901 aacaaacaaa caaaaattgc tcaaaaaaga ggagaaaagc tcaagaggaa aagctaagga
12961 ctggtaggaa aaagctttac tctttcatgc cattttattt ctttttgatt tttaaatcat
13021 tcattcaata gataccaccg tgtgacctat aattttgcaa atctgttacc tctgacatca
13081 agtgtaatta gcttttggag agtgggctga catcaagtgt aattagcttt tggagagtgg
13141 gttttgtcca ttattaataa ttaattaatt aacatcaaac acggcttctc atgctatttc
13201 tacctcactt tggttttggg gtgttcctga taattgtgca cacctgagtt cacagcttca
13261 ccacttgtcc attgcgttat tttcttttc ctttataatt ctttcttttt ccttcataat
13321 tttcaaaaga aaacccaaag ctctaaggta acaaattacc aaattacatg aagatttggt
13381 ttttgtcttg catttttttc ctttatgtga cgctggacct tttctttacc caaggatttt
13441 taaaactcag atttaaaaca aggggttact ttacatccta ctaagaagtt taagtaagta
13501 agtttcattc taaaatcaga ggtaaataga gtgcataaat aattttgttt taatcttttt
13561 gtttttcttt tagacacatt agctctggag tgagtctgtc ataatatttg aacaaaaatt
13621 gagagcttta ttgctgcatt ttaagcataa ttaatttgga cattatttcg tgttgtgttc
13681 tttataacca ccgagtatta aactgtaaat cataatgtaa ctgaagcata aacatcacat
13741 ggcatgtttt gtcattgttt tcaggtactg agttcttact tgagtatcat aatatattgt
13801 gttttaacac caacactgta acatttacga attatttttt taaacttcag ttttactgca
13861 ttttcacaac atatcagact tcaccaaata tatgccttac tattgtatta tagtactgct
13921 ttactgtgta tctcaataaa gcacgcagtt atgttac
```

FIGURE 11

Query=Human Dystrophin 1220-9328; Sbjt=Mouse Dystrophin 1238-9319

```
Query: 1220  gaagtaaacctggaccgtta 1239
             ||||||||  |||||  ||||
Sbjct: 1238  gaagtaaatctggatagtta 1257

Query: 1240  tcaaacagctttagaagaagtattatcgtggcttctttctgctgaggacacattgcaagc 1299
              |||||  |||||||||||||||  | ||  ||||||||||||||  |||||  |||||||  |||
Sbjct: 1258  ccaaactgctttagaagaagtactttcatggcttctttctgccgaggatacattgcgagc 1317

Query: 1300  acaaggagagatttctaatgatgtggaagtggtgaaagaccagtttcatactcatgaggg 1359
              ||||||||||||||  ||||||||  ||||  ||||||||  |||||||||  ||||||||||
Sbjct: 1318  acaaggagagatttcaaatgatgttgaagaagtgaaagaacagtttcatgctcatgaggg 1377

Query: 1360  gtacatgatggatttgacagcccatcagggccgggttggtaatattctacaattgggaag 1419
               | ||||||||||  |||||  |||||  ||  |  |||||||||  |||||||  ||  |||||
Sbjct: 1378  attcatgatggatctgacatctcatcaaggacttgttggtaatgttctacagttaggaag 1437

Query: 1420  taagctgattggaacaggaaaattatcagaagatgaagaaactgaagtacaagagcagat 1479
              | | ||  ||||||  |||  ||||||||||||||||||||||  |||||||  |||||  ||  ||
Sbjct: 1438  tcaactagttggaaaagggaaattatcagaagatgaagaagctgaagtgcaagaacaaat 1497

Query: 1480  gaatctcctaaattcaagatgggaatgcctcagggtagctagcatggaaaaacaaagcaa 1539
              ||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Sbjct: 1498  gaatctcctaaattcaagatgggaatgtctcagggtagctagcatggaaaaacaaagcaa 1557

Query: 1540  tttacat agagttttaatggatctccagaatcagaaactgaaagagttgaatgactggct 1599
               |||||  |  ||||  ||||||||||||||||||||||  |  |||||  |  ||||||||  |
Sbjct: 1558  attacac aaagttctaatggatctccagaatcagaaattaaaagaactagatgactggtt 1617

Query: 1600  aacaaaaacagaagaaagaacaaggaaaatggaggaagagcctcttggacctgatcttga 1659
              |||||||||  |||||  |||||  |  ||||||||||||||||||  ||||||||||||||||
Sbjct: 1618  aacaaaaactgaagagagaactaagaaaatggaggaagagccctttggacctgatcttga 1677

Query: 1660  agacctaaaacgccaagtacaacaacataaggtgcttcaagaagatctagaacaagaaca 1719
              |||  ||||||  ||||||||||||||||||||||||||||||||||||||||||||  ||  ||
Sbjct: 1678  agatctaaaatgccaagtacaacaacataaggtgcttcaagaagatctagaacaggagca 1737

Query: 1720  agtcagggtcaattctctcactcacatggtggtggtagttgatgaatctagtggagatca 1779
              ||||||||||  ||  ||||||||||||||  ||  ||  ||||||||||  ||  ||  |||||
Sbjct: 1738  ggtcagggtcaactcgctcactcacatggtagtagtggttgatgaatccagcggtgatca 1797

Query: 1780  cgcaactgctgctttggaagaacaacttaaggtattgggagatcgatgggcaaacatctg 1839
              |||||  ||||||||||||||||||||||||||  ||||||||||||||||||  |||||
Sbjct: 1798  tgcaacagctgctttggaagaacaacttaaggtactgggagatcgatgggcaaatatctg 1857

Query: 1840  tagatggacagaagaccgctgggttcttttacaagacatc cttctcaaatggcaacgtct 1899
              ||||||||  ||||||||||||  ||  ||||||||||  ||  |||||  ||||||||  |  |  |
Sbjct: 1858  cagatggactgaagaccgctggattgttttacaagatatt cttctaaaatggcagcattt 1917

Query: 1900  tactgaagaacagtgccttttttagtgcatggctttcagaaaaagaagatgcagtgaacaa 1959
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||||  ||
Sbjct: 1918  tactgaagaacagtgccttttttagtacatggctttcagaaaaagaagatgcaatgaagaa 1977

Query: 1960  gattcacacaactggctttaaagatcaaaatgaaatgttatcaagtcttcaaaaactggc 2019
              |||||  ||||  ||||||||||||||||||||||||  |  ||||||||||  |||  |  |
Sbjct: 1978  cattcagacaagtggctttaaagatcaaaatgaaatgatgtcaagtcttcacaaaatatc 2037

Query: 2020  cgttttaaaagcggatctagaaaagaaaaagcaatccatgggcaaactgtattcactcaa 2079
                 |||||||  ||||||||||||||||||  |  ||||||  |||||  |||||||||
Sbjct: 2038  tactttaaaaatagatctagaaaagaaaaagccaaccatggaaaaactaagttcactcaa 2097

Query: 2080  acaagatcttctttcaacactgaagaataagtcagtgacccagaagacggaagcatggct 2139
              ||||||||  |||||  |||||||  ||||||||||||||  ||  ||||  ||||   |||  |
Sbjct: 2098  tcaagatctactttcggcactgaaaaataagtcagtgactcaaaagatggaaatctggat 2157

Query: 2140  ggataactttgcccggtgttgggataatttagtccaaaaacttgaaaagagtacagcaca 2199
              |||  ||||||||  |  |||||||  ||||||  ||||||||||||||||||||  |||||||
Sbjct: 2158  ggaaaactttgcacaacgttgggacaatttaacccaaaaacttgaaaagagttcagcaca 2217
```

FIGURE 11 (cont.

```
Query: 2200  gatttcacaggct gtcaccaccactcagccatcactaacacagacaactgtaatggaaac 2259
Sbjct: 2218  aatttcacaggct gtcaccaccactcaaccatccctaacacagacaactgtaatggaaac 2277

Query: 2260  agtaactacggtgaccacaagggaacagatcctggtaaagcatgctcaagaggaacttcc 2319
Sbjct: 2278  ggtaactatggtgaccacaagggaacaaatcatggtaaaacatgcccaagaggaacttcc 2337

Query: 2320  accaccacctccccaaaagaagaggcagattactgtggat tctgaaattaggaaaaggtt 2379
Sbjct: 2338  accaccacctcctcaaaagaagaggcagataactgtggat tctgaactcaggaaaaggtt 2397

Query: 2380  ggatgttgatataactgaacttcacagctggattactcgctcagaagctgtgttgcagag 2439
Sbjct: 2398  ggatgtcgatataactgaacttcacagttggattactcgttcagaagctgtattacagag 2457

Query: 2440  tcctgaatttgcaatctttcggaaggaaggcaacttctcagacttaaaagaaaaagtcaa 2499
Sbjct: 2458  ttctgaatttgcagtctatcgaaaagaaggcaacatctcagacttgcaagaaaaagtcaa 2517

Query: 2500  tgccatagagcgagaaaaagctgagaagttcagaaaactgcaagatgccagcagatcagc 2559
Sbjct: 2518  tgccatagcacgagaaaaagcagagaagttcagaaaactgcaagatgccagcagatcagc 2577

Query: 2560  tcaggccctggtggaacagatggtgaatgagggtgttaatgcagatagcatcaaacaagc 2619
Sbjct: 2578  tcaggccctggtggaacagatggcaaatgagggtgttaatgctgaaagtatcagacaagc 2637

Query: 2620  ctcagaacaactgaacagccggtggatcgaattctgccagttgctaagtgagagacttaa 2679
Sbjct: 2638  ttcagaacaactgaacagccggtggacagaattctgccaattgctgagtgagagagttaa 2697

Query: 2680  ctggctggagtat cagaacaacatcatcgctttctataatcagctacaacaattggagca 2739
Sbjct: 2698  ctggctagagtat caaaccaacatcattacctttatataatcagctacaacaattggaaca 2757

Query: 2740  gatgacaactactgctgaaaactggttgaaaatccaacccaccaccccatcagagccaac 2799
Sbjct: 2758  gatgacaactactgccgaaaacttgttgaaaacccagtctaccaccctatcagagccaac 2817

Query: 2800  agcaattaaaagtcagttaaaaatttgtaaggatgaagtcaaccggctatcaggtcttca 2859
Sbjct: 2818  agcaattaaaagccagttaaaaatttgtaaggatgaagtcaacagattgtcagctcttca 2877

Query: 2860  acctcaaattgaacgattaaaaattcaaagcatagccctgaaagagaaaggacaaggacc 2919
Sbjct: 2878  gcctcaaattgagcaattaaaaattcagagtctacaactgaaagaaaagggacaggggcc 2937

Query: 2920  catgttcctggatgcagactttgtggcctttacaaatcattttaagcaagtcttttctga 2979
Sbjct: 2938  aatgtttctggatgcagactttgtggcctttactaatcattttaaccacatctttgatgg 2997

Query: 2980  tgtgcaggccagagagaaagagctacagacaatttttgac actttgccaccaatgcgcta 3039
Sbjct: 2998  tgtgagggccaaagagaaagagctacagacaatttttgac actttaccaccaatgcgcta 3057

Query: 3040  tcaggagaccatgagtgccatcaggacatgggtccagcagtcagaaaccaaactctccat 3099
Sbjct: 3058  tcaggagacaatgagtagcatcaggacgtggatccagcagtcagaaagcaaactctctgt 3117

Query: 3100  acctcaacttagtgtcaccgactatgaaatcatggagcagagactcggggaattgcaggc 3159
Sbjct: 3118  accttatcttagtgttactgaatatgaaataatggaggagagactcgggaaattacaggc 3177

Query: 3160  tttacaaagttctctgcaagagcaacaaagtggcctatactatctcagcaccactgtgaa 3219
Sbjct: 3178  tctgcaaagttctttgaaagagcaacaaaatggcttcaactatctgagtgacactgtgaa 3237
```

FIGURE 11 (cont.

```
Query: 3220  agagatgtcgaagaaagcgccctctgaaattagccggaaatatcaatcagaatttgaaga 3279
Sbjct: 3238  ggagatggccaagaaagcaccttcagaaatatgccagaaatatctgtcagaatttgaaga 3297

Query: 3280  aattgagggacgctggaagaagctctcctcccagctggttgagcattgtcaaaagctaga 3339
Sbjct: 3298  gattgaggggcactggaagaaactttcctcccagttggtggaaagctgccaaaagctaga 3357

Query: 3340  ggagcaa atgaataaactccgaaaaattcagaatcacatacaaaccctgaagaaatggat 3399
Sbjct: 3358  agaacat atgaataaacttcgaaaatttcagaatcacataaaaaccttacagaaatggat 3417

Query: 3400  ggctgaagttgatgtttttctgaaggaggaatggcctgcccttggggattcagaaattct 3459
Sbjct: 3418  ggctgaagttgatgttttcctgaaagaggaatggcctgccctgggggatgctgaaatcct 3477

Query: 3460  aaaaaagcagctgaaacagtgcagacttttagtcagtgatattcagacaattcagcccag 3519
Sbjct: 3478  gaaaaaacagctcaaacaatgcagacttttagttggtgatattcaaacaattcagcccag 3537

Query: 3520  tctaaacagtgtcaatgaaggtgggcagaagataaagaatgaagcagagccagagtttgc 3579
Sbjct: 3538  tttaaatagtgttaatgaaggtgggcagaagataaagagtgaagctgaacttgagtttgc 3597

Query: 3580  ttcgagacttgagacagaactcaaagaacttaacactcagtgggatcacatgtgccaaca 3639
Sbjct: 3598  atccagactggagacagaacttagagagcttaacactcagtgggatcacatatgccgcca 3657

Query: 3640  ggtctatgccagaaaggaggccttgaagggaggt ttggagaaaactgtaagcctccagaa 3699
Sbjct: 3658  ggtctacaccagaaaggaagccttaaaggcaggt ttggataaaaccgtaagcctccaaaa 3717

Query: 3700  agatctatcagagatgcacgaatggatgacacaagctgaagaagagtatcttgagagaga 3759
Sbjct: 3718  agatctatcagagatgcatgagtggatgacacaagctgaagaagaatatctagagagaga 3777

Query: 3760  ttttgaatataaaactccagatgaattacagaaagcagttgaagagatgaagagagctaa 3819
Sbjct: 3778  ttttgaatataaaactccagatgaattacagactgctgttgaagaaatgaagagagctaa 3837

Query: 3820  agaagaggcccaacaaaaagaagcgaaagtgaaactccttactgagtctgtaaatagtgt 3879
Sbjct: 3838  agaagaggcactacaaaaagaaactaaagtgaaactccttactgagactgtaaatagtgt 3897

Query: 3880  catagctcaagctccacctgtagcacaagaggccttaaaaaaggaacttgaaactctaac 3939
Sbjct: 3898  aatagctcacgctccaccctcagcacaagaggccttaaaaaaggaacttgaaactctgac 3957

Query: 3940  caccaactaccagtggctctgcactaggctgaatgggaaatgcaagactttggaagaagt 3999
Sbjct: 3958  caccaactaccaatggctgtgcaccaggctgaatggaaaatgcaaaactttggaagaagt 4017

Query: 4000  t tgggcatgttggcatgagttattgtcatacttggagaaagcaaacaagtggctaaatga 4059
Sbjct: 4018  t tgggcatgttggcatgagttattgtcatatttagagaaagcaaacaagtggctcaatga 4077

Query: 4060  agtagaatttaaacttaaaaccactgaaaacattcctggcggagctgaggaaatctctga 4119
Sbjct: 4078  agtagaattgaaacttaaaaccatggaaaatgttcctgcaggacctgaggaaatcactga 4137

Query: 4120  ggtgctagattcacttgaaaatttgatgcgacattcagaggataacccaaatcagattcg 4179
Sbjct: 4138  agtgctagaatctcttgaaaatctgatgcatcattcagaggagaacccaaatcagattcg 4197

Query: 4180  catattggcacagaccctaacagatggcggagtcatggatgagctaatcaatgaggaact 4239
Sbjct: 4198  tctattggcacagactcttacagatggaggagtcatggatgaactgatcaatgaggagct 4257
```

FIGURE 11 (cont.

```
Query: 4240  tgagacatttaattctcgttggagggaactacatgaagaggctgtaaggaggcaaaagtt 4299
             |||||| |||||||||||||||||||||||||||||||||||||| ||||  ||||||||
Sbjct: 4258  tgagacgtttaattctcgttggagggaactacatgaagaggctgtgaggaaacaaaagtt 4317

Query: 4300  gcttgaacagagc atccagtctgcccaggagactgaaaaatccttacacttaatccagga 4359
             ||||||||||||   ||||||||||||||||| | |||||| ||||| |||||||| |||||
Sbjct: 4318  gcttgaacagagt atccagtctgcccaggaaattgaaaagtccttgcacttaattcagga 4377

Query: 4360  gtccctcacattcattgacaagcagttggcagcttatattgcagacaaggtggacgcagc 4419
             ||| ||   ||||||||||||||||||||||||||||||  | ||||||||||| |||||
Sbjct: 4378  gtcgcttgaattcattgacaagcagttggcagcttatatcactgacaaggtggatgcagc 4437

Query: 4420  tcaaatgcctcaggaagcccagaaaatccaatctgatttgacaagtcatgagatcagttt 4479
             ||||||||||||||||||||||||||||||||| |||||||||||||||||||| |||||
Sbjct: 4438  tcaaatgcctcaggaagcccagaaaatccaatcagatttgacaagtcatgagataagttt 4497

Query: 4480  agaagaaatgaagaaacataatcaggggaaggaggctgcccaaagagtcctgtctcagat 4539
             ||||||||||||||||||||| ||||||||||| ||   |||||| || || || || ||
Sbjct: 4498  agaagaaatgaagaaacataaccaggggaaggatgccaaccaaagggttctttcacaaat 4557

Query: 4540  tgatgttgcacagaaaaaattacaagatgtctccatgaagtttcgatta ttccagaaacc 4599
             ||||||||||||||||||||||||||||||||||||||| ||||||||| ||||| |||||
Sbjct: 4558  tgatgttgcacagaaaaaattacaagatgtctccatgaaatttcgatta ttccaaaaacc 4617

Query: 4600  agccaattttgagctgcgtctacaagaaagtaagatgatttagatgaagtgaagatgca 4659
             |||||||||||| |  |||||| | ||||||||||||||||||||||||| ||||||||
Sbjct: 4618  agccaattttgaacaacgtctagaggaaagtaagatgatttagatgaagtcaagatgca 4677

Query: 4660  cttgcctgcattggaaacaaagagtgtggaacaggaagtagtacagtcacagctaaatca 4719
              |||||||||||||||||| |||||||| |||||||||||| | |||||||| |||| |||
Sbjct: 4678  tttgcctgcattggaaaccaagagtgttgaacaggaagtaattcagtcacaactaagtca 4737

Query: 4720  ttgtgaacttgtataaaagtctgagtgaagtgaagtctgaagtggaaatggtgataaa 4779
             ||||||||||||||||||| ||||||||||| ||||||||||||||||||||||| ||
Sbjct: 4738  ttgtgaacttgtataaaagcctgagtgaagtcaagtctgaagtggaaatggtgattaa 4797

Query: 4780  gactggacgtcagattgtacagaaaaagcagacggaaaatcccaaagaacttgatgaaag 4839
              || |||||||| |||||||||||||||||||| |||||||||||||| ||||||||| |
Sbjct: 4798  aaccggacgtcaaattgtacagaaaaagcagacagaaaatcccaaagagcttgatgaacg 4857

Query: 4840  agtaacagctttgaaattgcattataatgagctgggagcaaaggtaacagaaagaaagca 4899
             |||||||||||||||||||||||| |||||| |||| || ||||||||||| ||||||||
Sbjct: 4858  agtaacagctttgaaattgcattacaatgagttgggtgcgaaggtaacagagagaaagca 4917

Query: 4900  acagttggagaaatgc ttgaaattgtcccgtaagatgcgaaaggaaatgaatgtcttgac 4959
             |||||||||||||||| ||||| ||||||||||||||| ||||||||||||||||||| ||
Sbjct: 4918  acagttggagaaatgc ttgaagttgtcccgtaagatgagaaaggaaatgaatgtcttaac 4977

Query: 4960  agaatggctggcagctacagatatggaattgacaaagagatcagcagttgaaggaatgcc 5019
             ||||||||||||||| |||||||  |||||||| ||||||||||||||||||||||||||
Sbjct: 4978  agaatggctggcagcaacagatacagaattgacgaagagatcagcagttgaaggaatgcc 5037

Query: 5020  tagtaatttggattctgaagttgcctggggaaaggctactcaaaaagagattgagaaaca 5079
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 5038  aagtaatttggattctgaagttgcctggggaaaggctactcaaaaagagattgagaaaca 5097

Query: 5080  gaaggtgcacctgaagagtatcacagaggtaggagaggccttgaaaacagttttgggcaa 5139
             |||||  ||| |||||||| | |||||  |||||||| | |||||||  || ||||||||
Sbjct: 5098  gaaggctcacttgaagagtgttacagaattaggagagtctttgaaaatggtgttgggcaa 5157

Query: 5140  gaaggagacgttggtggaagataaactcagtcttctgaatagtaactggatagctgtcac 5199
             ||| || || ||||| ||||||||||| ||||||||||| ||||||||||||||||||||
Sbjct: 5158  gaaagaaaccttggtagaagataaactgagtcttctgaacagtaactggatagctgtcac 5217

Query: 5200  ctcccgagcagaagagtggttaaatcttttgttggaatac cagaaacacatggaaacttt 5259
             |||| ||| |||||| ||| |||||||||||||||||||| ||||||||||||||||| ||
Sbjct: 5218  ctccagagtagaagaatggctaaatcttttgttggaatac cagaaacacatggaaacctt 5277
```


24/6

FIGURE 11 (cont.

```
Query: 5260  tgaccagaatgtggaccacatcacaaagtggatcattcaggctgacacacttttggatga 5319
             ||| |||||  | || || ||||||||||||||||||| || ||   |||||| |||||
Sbjct: 5278  tgatcagaacatagaacaaatcacaaagtggatcattcatgcagatgaacttttagatga 5337

Query: 5320  atcagagaaaaagaaaccccagcaaaaagaagacgtgcttaagcgtttaaaggcagaact 5379
              || || || |||||||| || ||||| |||||| | |||||||||||||||||| ||| |
Sbjct: 5338  gtctgaaaagaagaaaccacaacaaaaggaagacattcttaagcgtttaaaggctgaaat 5397

Query: 5380  gaatgacatacgcccaaaggtggactctacacgtgaccaagcagcaaacttgatggcaaa 5439
              ||||||||| |||||||||||||||| |||||||||||||||||| ||||||||||
Sbjct: 5398  gaatgacatgcgcccaaaggtggactccacacgtgaccaagcagcaaaattgatggcaaa 5457

Query: 5440  ccgcggtgaccactgcaggaaattagtagagccccaaatctcagagctcaaccatcgatt 5499
              |||||||||||||||||||||| |||||||||||||||||| |||||||||| ||||||
Sbjct: 5458  ccgcggtgaccactgcaggaaagtagtagagccccaaatctctgagctcaaccgtcgatt 5517

Query: 5500  tgcagccatttcacacagaattaagactggaaaggcctccattcctttgaag gaattgga 5559
              |||||| ||||| |||||||||||||||||||||||||||||||||||||||| ||||||||
Sbjct: 5518  tgcagctatttctcacagaattaagactggaaaggcctccattcctttgaag gaattgga 5577

Query: 5560  gcagtttaactcagatatacaaaaattgcttgaaccactggaggctgaaattcagcaggg 5619
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 5578  gcagtttaactcagatatacaaaaattgcttgaaccactggaggctgaaattcagcaggg 5637

Query: 5620  ggtgaatctgaaagaggaagacttcaataaagatatgaatgaagacaatgagggtactgt 5679
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Sbjct: 5638  ggtgaatctgaaagaggaagacttcaataaagatatgagtgaagacaatgagggtactgt 5697

Query: 5680  aaaagaattgttgcaaagaggagacaacttacaacaaagaatcacagatgagagaaagag 5739
              ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Sbjct: 5698  aaatgaattgttgcaaagaggagacaacttacaacaaagaatcacagatgagagaaagcg 5757

Query: 5740  agaggaaataaagataaaacagcagctgttacagacaaaacataatgctctcaaggattt 5799
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 5758  agaggaaataaagataaaacagcagctgttacagacaaaacataatgctctcaaggattt 5817

Query: 5800  gaggtctcaaagaagaaaaaaggctctagaaatt tctcatcagtggtatcagtacaagag 5859
              |||||||||||||||||||||| ||||||||| ||||| ||||||||||||||||||||
Sbjct: 5818  gaggtctcaaagaagaaaaaaggccctagaaatt tctcaccagtggtatcagtacaagag 5877

Query: 5860  gcaggctgatgatctcctgaaatgcttggatgacattgnnnnnnnnttagccagcctacc 5919
              |||||||||||||||||||||||||||||||||| ||||        ||||||||||||||
Sbjct: 5878  gcaggctgatgatctcctgaaatgcttggatgaaattgaaaaaaaattagccagcctacc 5937

Query: 5920  tgagcccagagatgaaaggaaaataaaggaaattgatcgggaattgcagaagaagaaaga 5979
              ||| |||||||||||||| ||| |||||||||||||||| ||||||||||||||||||||
Sbjct: 5938  tgaacccagagatgaaagaaaattaaaggaaattgatcgtgaattgcagaagaagaaaga 5997

Query: 5980  ggagctgaatgcagtgcgtaggcaagctgagggcttgtctgaggatggggccgcaatggc 6039
              |||||||||||||||||| ||||||||||||||||||||||||| ||||||||||||||||
Sbjct: 5998  ggagctgaatgcagtgcgcaggcaagctgagggcttgtctgagaatggggccgcaatggc 6057

Query: 6040  agtggagccaactcagatccagctcagcaagcgctggcgggaaattgagagcaaatttgc 6099
              |||||||||||||||||||||||||||||||||||||||||||| |||||||||||| |||||
Sbjct: 6058  agtggagccaactcagatccagctcagcaagcgctggcggcaaattgagagcaattttgc 6117

Query: 6100  tcagtttcgaagactcaactttgcacaa attcacactgtccgtgaagaaacgatgatggt 6159
              |||||||||||||||||||||||||| ||||||||| ||| ||||||||| ||| | ||
Sbjct: 6118  tcagtttcgaagactcaactttgcacaa attcacactctccatgaagaaactatggtagt 6177

Query: 6160  gatgactgaagacatgcctttggaaatttcttatgtgccttctacttatttgactgaaat 6219
              || ||||||||| |||||||||||  |||||||||||||||||||||||||||| || ||
Sbjct: 6178  gacgactgaagatatgcctttggatgtttcttatgtgccttctacttatttgaccgagat 6237

Query: 6220  cactcatgtctcacaagccctattagaagtggaacaacttctcaatgctcctgacctctg 6279
              || |||| ||| |||||| || | |||||| || || ||||| ||| ||||||| ||||||
Sbjct: 6238  cagtcatatcttacaagctctttcagaagttgatcatcttctaaatactcctgaactctg 6297
```

FIGURE 11 (cont.

```
Query: 6280  tgctaaggactttgaagatctctttaagcaagaggagtctctgaagaatataaaagatag 6339
Sbjct: 6298  tgctaaagatttgaagatctttttaagcaagaggagtctcttaagaatataaaagacaa 6357

Query: 6340  tctacaacaaagctcaggtcggattgacattattcatagcaagaagacagcagcattgca 6399
Sbjct: 6358  tttgcaacaaatctcaggtcggattgatattattcacaagaagaagacagcagccttgca 6417

Query: 6400  aagtgcaacgcctgtggaaagggtgaagctacaggaagctctctcccagcttgatttcca 6459
Sbjct: 6418  aagtgccacctccatggaaaaggtgaaagtacaggaagccgtggcacagatggatttcca 6477

Query: 6460  atgggaaaaagttaacaaaatgtacaaggaccgacaagggcgatttgacagatct gttga 6519
Sbjct: 6478  ggggaaaaacttcatagaatgtacaaggaacgacaagggcgattcgacagatca gttga 6537

Query: 6520  gaaatggcggcgttttcattatgatataaagatatttaatcagtggctaacagaagctga 6579
Sbjct: 6538  aaaatggcgacactttcattatgatatgaaggtatttaatcaatggctgaatgaagttga 6597

Query: 6580  acagtttctcagaaagacacaaattcctgagaattgggaacatgctaaatacaaatggta 6639
Sbjct: 6598  acagttttcaaaaagacacaaaatcctgaaaactgggaacatgctaaatacaaatggta 6657

Query: 6640  tcttaaggaactccaggatggcattgggcagcggcaaactgttgtcagaacattgaatgc 6699
Sbjct: 6658  tcttaaggaactccaggatggcattgggcagcgtcaagctgttgtcagaacactgaatgc 6717

Query: 6700  aactggggaagaaataattcagcaatcctcaaaaacagatgccagtattctacaggaaaa 6759
Sbjct: 6718  aactggggaagaaataattcaacagtcttcaaaaacagatgtcaatattctacaagaaaa 6777

Query: 6760  attgggaagcctgaatctgcggtggcaggaggtctgcaaacagctgtcagacagaaaaaa 6819
Sbjct: 6778  attaggaagcttgagtctgcggtggcacgacatctgcaaagagctggcagaaaggagaaa 6837

Query: 6820  gaggctagaagaacaa aagaatatcttgtcagaatttcaaagagatttaaatgaatttgt 6879
Sbjct: 6838  gaggattgaagaacaa aagaatgtcttgtcagaatttcaaagagatttaaatgaatttgt 6897

Query: 6880  tttatggttggaggaagcagataacattgctagtatcccacttgaacctggaaaagagca 6939
Sbjct: 6898  tttgtggctggaagaagcagataacattgctattactccact------tggagatgagca 6951

Query: 6940  gcaactaaaagaaaagcttgagcaagtcaagttactggtggaagagttgcccctgcgcca 6999
Sbjct: 6952  gcagctaaaagaacaacttgaacaagtcaagttactggcagaagagttgcccctgcgcca 7011

Query: 7000  gggaattctcaaacaattaaatgaaactggaggacccgtgcttgtaagtgctcccataag 7059
Sbjct: 7012  gggaattctaaaacaattaaatgaaacaggaggagcagtacttgtaagtgctcccataag 7071

Query: 7060  cccagaagagcaagataaacttgaaaataagctcaagcagacaaatctccagtggataaa 7119
Sbjct: 7072  gccagaagagcaagataaacttgaaaagaagctcaaacagacaaatctccagtggataaa 7131

Query: 7120  ggtttccagagctttacctgagaaacaaggagaaattgaagctcaaataaaagaccttgg 7179
Sbjct: 7132  ggtctccagagctttacctgagaaacaaggagagcttgaggttcacttaaaagattttag 7191

Query: 7180  gcagctt gnnnnnnngcttgaagaccttgaagagcagttaaatcatctgctgctgtggtt 7239
Sbjct: 7192  gcag--- ------------------cttgaagagcagctggatcacctgcttctgtggct 7230

Query: 7240  atctcctattaggaatcagttggaaatttataaccaaccaaaccaagaaggaccatttga 7299
Sbjct: 7231  ctctcctattagaaaccagttggaaatttataaccaaccaagtcaggcaggaccgtttga 7290
```

26/66

FIGURE 11 (cont.

```
Query: 7300  cgttcaggaaactgaaatagcagttcaagctaaacaaccggatgtggaagagattttgtc 7359
Sbjct: 7291  cataaaggagattgaagtaacagttcacggtaaacaagcggatgtggaaaggcttttgtc 7350

Query: 7360  taaagggcagcatttgtacaaggaaaaaccagccactcagccagtgaagaggaagttaga 7419
Sbjct: 7351  gaaagggcagcatttgtataaggaaaaaccaagcactcagccagtgaagaggaagttaga 7410

Query: 7420  agatctgagctctgagtggaaggcggtaaaccgtttacttcaagagctgagggcaaagca 7479
Sbjct: 7411  agatctgaggtctgagtgggaggctgtaaaccatttacttcgggagctgaggacaaagca 7470

Query: 7480  gcctgaccta gctcctggactgaccactattggagcctctcctactcagactgttactct 7539
Sbjct: 7471  gcctgaccgt gcccctggactgagcactactggagcctctgccagtcagactgttactct 7530

Query: 7540  ggtgacacaacctgtggttactaaggaaactgccatctccaaactagaaatgccatcttc 7599
Sbjct: 7531  agtgacacaatctgtggttactaaggaaactgtcatctccaaactagaaatgccatcttc 7590

Query: 7600  cttgatgttggag gtacctgctctggcagatttcaaccgggcttggacagaacttaccga 7659
Sbjct: 7591  tttgctgttggag gtacctgcactggcagacttcaaccgagcttggacagaacttacaga 7650

Query: 7660  ctggctttctctgcttgatcaagttataaaatcacagagggtgatggtgggtgaccttga 7719
Sbjct: 7651  ctggctgtctctgcttgatcgagttataaaatcacagagagtgatggtgggtgatctgga 7710

Query: 7720  ggatatcaacgagatgatcatcaagcagaaggcaacaatgcaggatttggaacagaggcg 7779
Sbjct: 7711  agacatcaatgaaatgatcatcaaacagaaggcaacactgcaagatttggaacagagacg 7770

Query: 7780  tccccagttggaagaactcattaccgctgcccaaaatttgaaaaacaagaccagcaatca 7839
Sbjct: 7771  cccccaattggaagaactcattactgctgcccagaatttgaaaaacaaaaccagcaatca 7830

Query: 7840  agaggctagaacaatcattacggatcgaattgaaagaattcagaatcagtgggatgaagt 7899
Sbjct: 7831  agaagctagaacaatcattactgatcgaattgaaagaattcagattcagtgggatgaggt 7890

Query: 7900  acaagaacaccttcagaaccggaggcaacagttgaatgaaatg ttaaaggattcaacaca 7959
Sbjct: 7891  tcaagaacagctgcagaacaggagacaacagttgaatgaaatg ttaaaggattcaacaca 7950

Query: 7960  atggctggaagctaaggaagaagctgagcaggtcttaggacaggccagagccaagcttga 8019
Sbjct: 7951  atggctggaagctaaggaagaagccgaacaggtcataggacaggtcagaggcaagcttga 8010

Query: 8020  gtcatggaaggagggtccctatacagtagatgcaatccaaaagaaaatcacagaaaccaa 8079
Sbjct: 8011  ctcatggaaagaaggtcctcacacagtagatgcaatccaaaagaagatcacagaaaccaa 8070

Query: 8080  gcagttggccaaagacctccgccagtggcagacaaatgtagatgtggcaaatgacttggc 8139
Sbjct: 8071  gcagttggccaaagacctccgtcaacggcagataagtgtagacgtggcaaatgatttggc 8130

Query: 8140  cctgaaacttctccgggattattctgcagatgataccagaaaagtccacatgataacaga 8199
Sbjct: 8131  actgaaacttcttcgggactattctgctgatgataccagaaaagtacacatgataacaga 8190

Query: 8200  gaatatcaatgcctcttggagaagcattcataaaagggtgagtgagcgagaggctgcttt 8259
Sbjct: 8191  gaatatcaatacttcttggggaaacattcataaaagagtaagtgagcaagaggctgcttt 8250

Query: 8260  ggaagaaact catagattactgcaacagttccccctggacctggaaaagtttcttgcctg 8319
Sbjct: 8251  ggaagaaact catagattactgcagcagttccctctggacctggagaagtttctttcctg 8310
```

FIGURE 11 (cont.

```
Query: 8320  gcttacagaagctgaaacaactgccaatgtcctacaggatgctacccgtaaggaaaggct 8379
Sbjct: 8311  gattacggaagcagaaacaactgccaatgtcctacaggacgcttcccgtaaggagaagct 8370

Query: 8380  cctagaagactccaagggagtaaaagagctgatgaaacaatggcaagacctccaaggtga 8439
Sbjct: 8371  cctagaagactccaggggagtcagagagctgatgaaaccatggcaagatctccaaggaga 8430

Query: 8440  aattgaagctcacacagatgtttatcacaacctggatgaaaacagccaaaaaatcctgag 8499
Sbjct: 8431  aattgaaactcacacagatatctatcacaatcttgatgaaaatggccaaaaaatcctgag 8490

Query: 8500  atccctggaaggttccgatgatgcagtcctgttacaaagacgtttggataacatgaactt 8559
Sbjct: 8491  atccctggaaggttcggatgaagcacccctgttacaaagacgtttggataacatgaattt 8550

Query: 8560  caagtggagtgaacttcggaaaaagtctctcaacattaggtcccatttggaagccagt tc 8619
Sbjct: 8551  caagtggagtgaacttcagaaaaagtctctcaacattaggtcccatttggaagcaagt tc 8610

Query: 8620  tgaccagtggaagcgtctgcacctttctctgcaggaacttctggtgtggctacagctgaa 8679
Sbjct: 8611  tgaccagtggaagcgtttgcatctttctcttcaggaacttcttgtttggctacagctgaa 8670

Query: 8680  agatgatgaattaagccggcaggcacctattggaggcgactttccagcagttcagaagca 8739
Sbjct: 8671  agatgatgaactgagccgtcaggcacccatcggtggtgatttcccagcagttcagaagca 8730

Query: 8740  gaacgatgtacatagggccttcaagagggaattgaaaactaaagaacctgtaatcatgag 8799
Sbjct: 8731  gaatgatatacatagggccttcaagagggaattgaaaactaaagaacctgtaatcatgag 8790

Query: 8800  tactcttgagactgtacgaatatttctgacagagcagcctttggaaggactagagaaact 8859
Sbjct: 8791  tactctggagactgtgagaatatttctgacagagcagcctttggaaggactagagaaact 8850

Query: 8860  ctaccaggagcccagagagctgcctcctgaggagagagcccagaatgtcactcggcttct 8919
Sbjct: 8851  ctaccaggagcccagagaactgcctcctgaagaaagagctcagaatgtcactcggctcct 8910

Query: 8920  acgaaagcaggctgaggaggtcaatactgagtgggaaaaattgaacctgcactccgctga 8979
Sbjct: 8911  acgaaagcaggctgaagaggtcaacgctgaatgggacaaattgaacctgcgctcagctga 8970

Query: 8980  ctggcagagaaaaatagatgagacc cttgaaagactccaggaacttcaagaggccacgga 9039
Sbjct: 8971  ttggcagagaaaaatagatgaagct cttgaaagactccaggaacttcaggaagctgccga 9030

Query: 9040  tgagctggacctcaagctgcgccaagctgaggtgatcaagggatcctggcagcccgtggg 9099
Sbjct: 9031  tgaactggacctcaagttgcgccaagctgaggtgatcaagggatcctggcagccagtggg 9090

Query: 9100  cgatctcctcattgactctctccaagatcacctcgagaaagtcaaggcacttcgaggaga 9159
Sbjct: 9091  ggatctcctcattgactctctgcaagatcaccttgaaaaagtcaaggcacttcggggaga 9150

Query: 9160  aattgcgcctctgaaagagaacgtgagccacgtcaatgaccttgctcgccagcttaccac 9219
Sbjct: 9151  aattgcacctcttaaagagaatgtcaatcgtgtcaatgaccttgcacatcagctgaccac 9210

Query: 9220  tttgggcattcagctctcaccgtataacctcagcactctggaagacctgaacaccagatg 9279
Sbjct: 9211  actgggcattcagctctcaccttataacctcagcactttggaagatctgaataccagatg 9270

Query: 9280  gaagcttctgcaggtggccgtcgaggaccgagtcaggcagctgcatgaa
Sbjct: 9271  gaggcttctacaggtggctgtggaggaccgtgtcagacagctgcatgaa
```

FIGURE 12 (ΔR4-R23, SEQ ID NO:39)

GGGATTCCCTCACTTTCCCCCTACAGGACTCAGATCTGGGAGGCAATTACCTTCGGAGAAAAACGAATAGGA
AAAACTGAAGTGTTACTTTTTTAAAGCTGCTGAAGTTTGTTGGTTTCTCATTGTTTTTAAGCCTACTGGAG
CAATAAAGTTTGAAGAACTTTTACCAGGTTTTTTTATCGCTGCCTTGATATACACTTTTCAAAATGCTTTG
GTGGGAAGAAGTAGAGGACTGTTATGAAAGAGAAGATGTTCAAAAGAAAACATTCACAAAATGGGTAAATGC
ACAATTTTCTAAGTTTGGGAAGCAGCATATTGAGAACCTCTTCAGTGACCTACAGGATGGGAGGCGCCTCCT
AGACCTCCTCGAAGGCCTGACAGGGCAAAAACTGCCAAAAGAAAAAGGATCCACAAGAGTTCATGCCCTGAA
CAATGTCAACAAGGCACTGCGGGTTTTGCAGAACAATAATGTTGATTTAGTGAATATTGGAAGTACTGACAT
CGTAGATGGAAATCATAAACTGACTCTTGGTTTGATTTGGAATATAATCCTCCACTGGCAGGTCAAAAATGT
AATGAAAAATATCATGGCTGGATTGCAACAAACCAACAGTGAAAAGATTCTCCTGAGCTGGGTCCGACAATC
AACTCGTAATTATCCACAGGTTAATGTAATCAACTTCACCACCAGCTGGTCTGATGGCCTGGCTTTGAATGC
TCTCATCCATAGTCATAGGCCAGACCTATTTGACTGGAATAGTGTGGTTTGCCAGCAGTCAGCCACACAACG
ACTGGAACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGATGTTGA
TACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCAAGTTTTGCCTCAACAAGT
GAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCCACCTAAAGTGACTAAAGAAGAACATTTTCA
GTTACATCATCAAATGCACTATTCTCAACAGATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTC
CCCTAAGCCTCGATTCAAGAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAG
CCCATTTCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAGTGAAGT
AAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCA
AGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGACCAGTTTCATACTCATGAGGGGTACATGAT
GGATTTGACAGCCCATCAGGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAA
ATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAG
GGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACTGAAAGA
GTTGAATGACTGGCTAACAAAAACAGAAGAAAGAACAAGGAAAATGGAGGAAGAGCCTCTTGGACCTGATCT
TGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAGAACAAGTCAGGGT
CAATTCTCTCACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGA
ACAACTTAAGGTATTGGGAGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACA
AGACATCCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTCAGAAAAAGA
AGATGCAGTGAACAAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTTATCAAGTCTTCAAAAACT
GGCCGTTTTAAAAGCGGATCTAGAAAAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAACAAGATCT
TCTTTCAACACTGAAGAATAAGTCAGTGACCCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTG
GGATAATTTAGTCCAAAAACTTGAAAAGAGTACAGCACAGATTTCACAGGCTGTCACCACCACTCAGCCATC
ACTAACACAGACAACTGTAATGGAAACAGTAACTACGGTGACCACAAGGGAACAGATCCTGGTAAAGCATGC
TCAAGAGGAACTTCCACCACCACCTCCCCAAAAGAAGAGGCAGATTACTGTGGATCTTGAAAGACTCCAGGA
ACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCC
CGTGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGC
GCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTC
ACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCG
AGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCA
GGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTG
CTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTA
TAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATG
TGATGCCTTGGACCAGCACAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTT
GACCACTATTTATGACCGCCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCCTCTCTGCGTGGATATGTG
TCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCTGTCTTTTAAAACTGG
CATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAAC
AGGATTTTGTGACCAGCGCAGGCTGGGCCTCCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGGTGA
AGTTGCATCCTTTGGGGGCAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTGCTAATAATAAGCC
AGAGATCGAAGCGGCCCTCTTCCTAGACTGGATGAGACTGGAACCCCAGTCCATGGTGTGGCTGCCCGTCCT
GCACAGAGTGGCTGCTGCAGAAACTGCCAAGCATCAGGCCAAATGTAACATCTGCAAAGAGTGTCCAATCAT
TGGATTCAGGTACAGGAGTCTAAAGCACTTTAATTATGACATCTGCCAAAGCTGCTTTTTTTCTGGTCGAGT
TGCAAAAGGCCATAAAATGCACTATCCCATGGTGGAATATTGCACTCCGACTACATCAGGAGAAGATGTTCG
AGACTTTGCCAAGGTACTAAAAAACAAATTTCGAACCAAAAGGTATTTTGCGAAGCATCCCCGAATGGGCTA
CCTGCCAGTGCAGACTGTCTTAGAGGGGGACAACATGGAAACGCCTGCCTCGTCCCCTCAGCTTTCACACGA

FIGURE 12 (cont.)

TGATACTCATTCACGCATTGAACATTATGCTAGCAGGCTAGCAGAAATGGAAAACAGCAATGGATCTTATCT
AAATGATAGCATCTCTCCTAATGAGAGCATAGATGATGAACATTTGTTAATCCAGCATTACTGCCAAAGTTT
GAACCAGGACTCCCCCCTGAGCCAGCCTCGTAGTCCTGCCCAGATCTTGATTTCCTTAGAGAGTGAGGAAAG
AGGGGAGCTAGAGAGAATCCTAGCAGATCTTGAGGAAGAAAACAGGAATCTGCAAGCAGAATATGACCGTCT
AAAGCAGCAGCACGAACATAAAGGCCTGTCCCCACTGCCGTCCCCTCCTGAAATGATGCCCACCTCTCCCCA
GAGTCCCCGGGATGCTGAGCTCATTGCTGAGGCCAAGCTACTGCGTCAACACAAAGGCCGCCTGGAAGCCAG
GATGCAAATCCTGGAAGACCACAATAAACAGCTGGAGTCACAGTTACACAGGCTAAGGCAGCTGCTGGAGCA
ACCCCAGGCAGAGGCCAAAGTGAATGGCACAACGGTGTCCTCTCCTTCTACCTCTCTACAGAGGTCCGACAG
CAGTCAGCCTATGCTGCTCCGAGTGGTTGGCAGTCAAACTTCGGACTCCATGGGTGAGGAAGATCTTCTCAG
TCCTCCCCAGGACACAAGCACAGGGTTAGAGGAGGTGATGGAGCAACTCAACAACTCCTTCCCTAGTTCAAG
AGGAAGAAATACCCCTGGAAAGCCAATGAGAGAGGACACAATGTAGGAAGTCTTTTCCACATGGCAGATGAT
TTGGGCAGAGCGATGGAGTCCTTAGTATCAGTCATGACAGATGAAGAAGGAGCAGAATAAATGTTTTACAAC
TCCTGATTCCCGCATGGTTTTTATAATATTCATACAACAAAGAGGATTAGACAGTAAGAGTTTACAAGAAAT
AAATCTATATTTTTGTGAAGGGTAGTGGTATTATACTGTAGATTTCAGTAGTTTCTAAGTCTGTTATTGTTT
TGTTAACAATGGCAGGTTTTACACGTCTATGCAATTGTACAAAAAAGTTATAAGAAAACTACATGTAAAATC
TTGATAGCTAAATAACTTGCCATTTCTTTATATGGAACGCATTTTGGGTTGTTTAAAAATTTATAACAGTTA
TAAAGAAAGATTGTAAACTAAAGTGTGCTTTATAAAAAAAAGTTGTTTATAAAAACCCCTAAAAACAAAACA
AACACACACACACACACATACACACACACACACAAAACTTTGAGGCAGCGCATTGTTTTGCATCCTTTTGGC
GTGATATCCATATGAAATTCATGGCTTTTTCTTTTTTTGCATATTAAAGATAAGACTTCCTCTACCACCACA
CCAAATGACTACTACACACTGCTCATTTGAGAACTGTCAGCTGAGTGGGGCAGGCTTGAGTTTTCATTTCAT
ATATCTATATGTCTATAAGTATATAAATACTATAGTTATATAGATAAAGAGATACGAATTTCTATAGACTGA
CTTTTTCCATTTTTTAAATGTTCATGTCACATCCTAATAGAAAGAAATTACTTCTAGTCAGTCATCCAGGCT
TACCTGCTTGGTCTAGA

FIGURE 13 (ΔR2-R21, SEQ ID NO:40)

GGGATTCCCTCACTTTCCCCCTACAGGACTCAGATCTGGGAGGCAATTACCTTCGGAGAAAAACGAATAGGA
AAAACTGAAGTGTTACTTTTTTTAAAGCTGCTGAAGTTTGTTGGTTTCTCATTGTTTTTAAGCCTACTGGAG
CAATAAAGTTTGAAGAACTTTTACCAGGTTTTTTTTATCGCTGCCTTGATATACACTTTTCAAAATGCTTTG
GTGGGAAGAAGTAGAGGACTGTTATGAAAGAGAAGATGTTCAAAAGAAAACATTCACAAAATGGGTAAATGC
ACAATTTTCTAAGTTTGGGAAGCAGCATATTGAGAACCTCTTCAGTGACCTACAGGATGGGAGGCGCCTCCT
AGACCTCCTCGAAGGCCTGACAGGGCAAAAACTGCCAAAAGAAAAAGGATCCACAAGAGTTCATGCCCTGAA
CAATGTCAACAAGGCACTGCGGGTTTTGCAGAACAATAATGTTGATTTAGTGAATATTGGAAGTACTGACAT
CGTAGATGGAAATCATAAACTGACTCTTGGTTTGATTTGGAATATAATCCTCCACTGGCAGGTCAAAAATGT
AATGAAAAATATCATGGCTGGATTGCAACAAACCAACAGTGAAAAGATTCTCCTGAGCTGGGTCCGACAATC
AACTCGTAATTATCCACAGGTTAATGTAATCAACTTCACCACCAGCTGGTCTGATGGCCTGGCTTTGAATGC
TCTCATCCATAGTCATAGGCCAGACCTATTTGACTGGAATAGTGTGGTTTGCCAGCAGTCAGCCACACAACG
ACTGGAACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGATGTTGA
TACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCAAGTTTTGCCTCAACAAGT
GAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCCACCTAAAGTGACTAAAGAAGAACATTTTCA
GTTACATCATCAAATGCACTATTCTCAACAGATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTC
CCCTAAGCCTCGATTCAAGAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAG
CCCATTTCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAGTGAAGT
AAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCA
AGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGACCAGTTTCATACTCATGAGGGGTACATGAT
GGATTTGACAGCCCATCAGGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAA
ATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAG
GGTAGCTAGCATGGAAAAACAAAGCAATTTACATCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAA
GTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCT
CCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCA
CACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGC
AGTCCTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACAT
TAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTG
GCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCA
GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGAC
TGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCC
TCCTGAGGAGAGAGCCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA
AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCCAGGAACTTCA
AGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG
CGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCT
GAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTA
TAACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG
GCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCC
CTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTGGGA
CCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGAC
TGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGC
CTTGGACCAGCACAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC
TATTTATGACCGCCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCCTCTCTGCGTGGATATGTGTCTGAA
CTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCTGTCTTTTAAAACTGGCATCAT
TTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATT
TTGTGACCAGCGCAGGCTGGGCCTCCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGC
ATCCTTTGGGGGCAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTGCTAATAATAAGCCAGAGAT
CGAAGCGGCCCTCTTCCTAGACTGGATGAGACTGGAACCCCAGTCCATGGTGTGGCTGCCCGTCCTGCACAG
AGTGGCTGCTGCAGAAACTGCCAAGCATCAGGCCAAATGTAACATCTGCAAAGAGTGTCCAATCATTGGATT
CAGGTACAGGAGTCTAAAGCACTTTAATTATGACATCTGCCAAAGCTGCTTTTTTTCTGGTCGAGTTGCAAA
AGGCCATAAAATGCACTATCCCATGGTGGAATATTGCACTCCGACTACATCAGGAGAAGATGTTCGAGACTT
TGCCAAGGTACTAAAAAACAAATTTCGAACCAAAAGGTATTTTGCGAAGCATCCCCGAATGGGCTACCTGCC
AGTGCAGACTGTCTTAGAGGGGGACAACATGGAAACGCCTGCCTCGTCCCCTCAGCTTTCACACGATGATAC
TCATTCACGCATTGAACATTATGCTAGCAGGCTAGCAGAAATGGAAAACAGCAATGGATCTTATCTAAATGA

FIGURE 13 (cont.)

TAGCATCTCTCCTAATGAGAGCATAGATGATGAACATTTGTTAATCCAGCATTACTGCCAAAGTTTGAACCA
GGACTCCCCCCTGAGCCAGCCTCGTAGTCCTGCCCAGATCTTGATTTCCTTAGAGAGTGAGGAAAGAGGGGA
GCTAGAGAGAATCCTAGCAGATCTTGAGGAAGAAAACAGGAATCTGCAAGCAGAATATGACCGTCTAAAGCA
GCAGCACGAACATAAAGGCCTGTCCCCACTGCCGTCCCCTCCTGAAATGATGCCCACCTCTCCCCAGAGTCC
CCGGGATGCTGAGCTCATTGCTGAGGCCAAGCTACTGCGTCAACACAAAGGCCGCCTGGAAGCCAGGATGCA
AATCCTGGAAGACCACAATAAACAGCTGGAGTCACAGTTACACAGGCTAAGGCAGCTGCTGGAGCAACCCCA
GGCAGAGGCCAAAGTGAATGGCACAACGGTGTCCTCTCCTTCTACCTCTCTACAGAGGTCCGACAGCAGTCA
GCCTATGCTGCTCCGAGTGGTTGGCAGTCAAACTTCGGACTCCATGGGTGAGGAAGATCTTCTCAGTCCTCC
CCAGGACACAAGCACAGGGTTAGAGGAGGTGATGGAGCAACTCAACAACTCCTTCCCTAGTTCAAGAGGAAG
AAATACCCCTGGAAAGCCAATGAGAGAGGACACAATGTAGGAAGTCTTTTCCACATGGCAGATGATTTGGGC
AGAGCGATGGAGTCCTTAGTATCAGTCATGACAGATGAAGAAGGAGCAGAATAAATGTTTTACAACTCCTGA
TTCCCGCATGGTTTTTATAATATTCATACAACAAAGAGGATTAGACAGTAAGAGTTTACAAGAAATAAATCT
ATATTTTTGTGAAGGGTAGTGGTATTATACTGTAGATTTCAGTAGTTTCTAAGTCTGTTATTGTTTTGTTAA
CAATGGCAGGTTTTACACGTCTATGCAATTGTACAAAAAAGTTATAAGAAAACTACATGTAAAATCTTGATA
GCTAAATAACTTGCCATTTCTTTATATGGAACGCATTTTGGGTTGTTTAAAAATTTATAACAGTTATAAAGA
AAGATTGTAAACTAAAGTGTGCTTTATAAAAAAAAAGTTGTTTATAAAAACCCCTAAAAACAAAACAAACACA
CACACACACACATACACACACACACAAAACTTTGAGGCAGCGCATTGTTTTGCATCCTTTTGGCGTGATA
TCCATATGAAATTCATGGCTTTTTCTTTTTTTGCATATTAAAGATAAGACTTCCTCTACCACCACACCAAAT
GACTACTACACACTGCTCATTTGAGAACTGTCAGCTGAGTGGGGCAGGCTTGAGTTTTCATTTCATATATCT
ATATGTCTATAAGTATATAAATACTATAGTTATATAGATAAAGAGATACGAATTTCTATAGACTGACTTTTT
CCATTTTTTAAATGTTCATGTCACATCCTAATAGAAAGAAATTACTTCTAGTCAGTCATCCAGGCTTACCTG
CTTGGTCTAGA

FIGURE 14 (ΔR2-R21+H3, SEQ ID NO:41)

GGGATTCCCTCACTTTCCCCCTACAGGACTCAGATCTGGGAGGCAATTACCTTCGGAGAAAAACGAATAGGA
AAAACTGAAGTGTTACTTTTTTTAAAGCTGCTGAAGTTTGTTGGTTTCTCATTGTTTTTAAGCCTACTGGAG
CAATAAAGTTTGAAGAACTTTTACCAGGTTTTTTTTATCGCTGCCTTGATATACACTTTTCAAAATGCTTTG
GTGGGAAGAAGTAGAGGACTGTTATGAAAGAGAAGATGTTCAAAAGAAAACATTCACAAAATGGGTAAATGC
ACAATTTTCTAAGTTTGGGAAGCAGCATATTGAGAACCTCTTCAGTGACCTACAGGATGGGAGGCGCCTCCT
AGACCTCCTCGAAGGCCTGACAGGGCAAAAACTGCCAAAAGAAAAAGGATCCACAAGAGTTCATGCCCTGAA
CAATGTCAACAAGGCACTGCGGGTTTTGCAGAACAATAATGTTGATTTAGTGAATATTGGAAGTACTGACAT
CGTAGATGGAAATCATAAACTGACTCTTGGTTTGATTTGGAATATAATCCTCCACTGGCAGGTCAAAAATGT
AATGAAAAATATCATGGCTGGATTGCAACAAACCAACAGTGAAAAGATTCTCCTGAGCTGGGTCCGACAATC
AACTCGTAATTATCCACAGGTTAATGTAATCAACTTCACCACCAGCTGGTCTGATGGCCTGGCTTTGAATGC
TCTCATCCATAGTCATAGGCCAGACCTATTTGACTGGAATAGTGTGGTTTGCCAGCAGTCAGCCACACAACG
ACTGGAACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGATGTTGA
TACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCAAGTTTTGCCTCAACAAGT
GAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCCACCTAAAGTGACTAAAGAAGAACATTTTCA
GTTACATCATCAAATGCACTATTCTCAACAGATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTC
CCCTAAGCCTCGATTCAAGAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAG
CCCATTTCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAGTGAAGT
AAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCA
AGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGACCAGTTTCATACTCATGAGGGGTACATGAT
GGATTTGACAGCCCATCAGGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAA
ATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAG
GGTAGCTAGCATGGAAAAACAAAGCAATTTACATGCTCCTGGACTGACCACTATTGGAGCCTCTCCTACTCA
GACTGTTACTCTGGTGACACAACCTGTGGTTACTAAGGAAACTGCCATCTCCAAACTAGAAATGCCATCTTC
CTTGATGTTGGAGCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGA
AGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGT
AAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCT
GGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTT
GGATAACATGAACTTCAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAG
TTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA
ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTT
CAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGA
GCAGCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAA
TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGC
TGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA
CCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTGACTCTCT
CCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGT
CAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA
CCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCACAG
GGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC
AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCT
CTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT
GCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCAGCACAACCTCAA
GCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTTATGACCGCCTGGAGCA
AGAGCACAACAATTTGGTCAACGTCCCTCTCTGCGTGGATATGTGTCTGAACTGGCTGCTGAATGTTTATGA
TACGGGACGAACAGGGAGGATCCGTGTCCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTT
GGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGG
CCTCCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGCAGTAACAT
TGAGCCAAGTGTCCGGAGCTGCTTCCAATTTGCTAATAATAAGCCAGAGATCGAAGCGGCCCTCTTCCTAGA
CTGGATGAGACTGGAACCCCAGTCCATGGTGTGGCTGCCCGTCCTGCACAGAGTGGCTGCTGCAGAAACTGC
CAAGCATCAGGCCAAATGTAACATCTGCAAAGAGTGTCCAATCATTGGATTCAGGTACAGGAGTCTAAAGCA
CTTTAATTATGACATCTGCCAAAGCTGCTTTTTTTCTGGTCGAGTTGCAAAAGGCCATAAAATGCACTATCC
CATGGTGGAATATTGCACTCCGACTACATCAGGAGAAGATGTTCGAGACTTTGCCAAGGTACTAAAAAACAA
ATTTCGAACCAAAAGGTATTTTGCGAAGCATCCCCGAATGGGCTACCTGCCAGTGCAGACTGTCTTAGAGGG

FIGURE 14 (cont.)

GGACAACATGGAAACGCCTGCCTCGTCCCCTCAGCTTTCACACGATGATACTCATTCACGCATTGAACATTA
TGCTAGCAGGCTAGCAGAAATGGAAAACAGCAATGGATCTTATCTAAATGATAGCATCTCTCCTAATGAGAG
CATAGATGATGAACATTTGTTAATCCAGCATTACTGCCAAAGTTTGAACCAGGACTCCCCCCTGAGCCAGCC
TCGTAGTCCTGCCCAGATCTTGATTTCCTTAGAGAGTGAGGAAAGAGGGGAGCTAGAGAGAATCCTAGCAGA
TCTTGAGGAAGAAAACAGGAATCTGCAAGCAGAATATGACCGTCTAAAGCAGCAGCACGAACATAAAGGCCT
GTCCCCACTGCCGTCCCCTCCTGAAATGATGCCCACCTCTCCCCAGAGTCCCCGGGATGCTGAGCTCATTGC
TGAGGCCAAGCTACTGCGTCAACACAAAGGCCGCCTGGAAGCCAGGATGCAAATCCTGGAAGACCACAATAA
ACAGCTGGAGTCACAGTTACACAGGCTAAGGCAGCTGCTGGAGCAACCCCAGGCAGAGGCCAAAGTGAATGG
CACAACGGTGTCCTCTCCTTCTACCTCTCTACAGAGGTCCGACAGCAGTCAGCCTATGCTGCTCCGAGTGGT
TGGCAGTCAAACTTCGGACTCCATGGGTGAGGAAGATCTTCTCAGTCCTCCCCAGGACACAAGCACAGGGTT
AGAGGAGGTGATGGAGCAACTCAACAACTCCTTCCCTAGTTCAAGAGGAAGAAATACCCCTGGAAAGCCAAT
GAGAGAGGACACAATGTAGGAAGTCTTTTCCACATGGCAGATGATTTGGGCAGAGCGATGGAGTCCTTAGTA
TCAGTCATGACAGATGAAGAAGGAGCAGAATAAATGTTTTACAACTCCTGATTCCCGCATGGTTTTTATAAT
ATTCATACAACAAAGAGGATTAGACAGTAAGAGTTTACAAGAAATAAATCTATATTTTTGTGAAGGGTAGTG
GTATTATACTGTAGATTTCAGTAGTTTCTAAGTCTGTTATTGTTTTGTTAACAATGGCAGGTTTTACACGTC
TATGCAATTGTACAAAAAAGTTATAAGAAAACTACATGTAAAATCTTGATAGCTAAATAACTTGCCATTTCT
TTATATGGAACGCATTTTGGGTTGTTTAAAAATTTATAACAGTTATAAAGAAAGATTGTAAACTAAAGTGTG
CTTTATAAAAAAAAGTTGTTTATAAAAACCCCTAAAAACAAAACAAACACACACACACACACATACACACAC
ACACACAAAACTTTGAGGCAGCGCATTGTTTTGCATCCTTTTGGCGTGATATCCATATGAAATTCATGGCTT
TTTCTTTTTTTGCATATTAAAGATAAGACTTCCTCTACCACCACACCAAATGACTACTACACACTGCTCATT
TGAGAACTGTCAGCTGAGTGGGGCAGGCTTGAGTTTTCATTTCATATATCTATATGTCTATAAGTATATAAA
TACTATAGTTATATAGATAAAGAGATACGAATTTCTATAGACTGACTTTTTCCATTTTTTAAATGTTCATGT
CACATCCTAATAGAAAGAAATTACTTCTAGTCAGTCATCCAGGCTTACCTGCTTGGTCTAGA

FIGURE 15 (ΔH2-R19, SEQ ID NO:42)

gggattccct cactttcccc ctacaggact cagatctggg aggcaattac cttcggagaa
aaacgaatag gaaaaactga agtgttactt tttttaaagc tgctgaagtt tgttggtttc
tcattgtttt taagcctact ggagcaataa agtttgaaga acttttacca ggtttttttt
atcgctgcct tgatatacac ttttcaaaat gctttggtgg gaagaagtag aggactgtta
tgaaagagaa gatgttcaaa agaaaacatt cacaaaatgg gtaaatgcac aattttctaa
gtttgggaag cagcatattg agaacctctt cagtgaccta caggatggga ggcgcctcct
agacctcctc gaaggcctga cagggcaaaa actgccaaaa gaaaaaggat ccacaagagt
tcatgccctg aacaatgtca acaaggcact gcgggttttg cagaacaata atgttgattt
agtgaatatt ggaagtactg acatcgtaga tggaaatcat aaactgactc ttggtttgat
ttggaatata atcctccact ggcaggtcaa aaatgtaatg aaaaatatca tggctggatt
gcaacaaacc aacagtgaaa agattctcct gagctgggtc cgacaatcaa ctcgtaatta
tccacaggtt aatgtaatca acttcaccac cagctggtct gatggcctgg ctttgaatgc
tctcatccat agtcataggc cagacctatt tgactggaat agtgtggttt gccagcagtc
agccacacaa cgactggaac atgcattcaa catcgccaga tatcaattag gcatagagaa
actactcgat cctgaagatg ttgataccac ctatccagat aagaagtcca tcttaatgta
catcacatca ctcttccaag ttttgcctca acaagtgagc attgaagcca tccaggaagt
ggaaatgttg ccaaggccac ctaaagtgac taaagaagaa cattttcagt tacatcatca
aatgcactat tctcaacaga tcacggtcag tctagcacag ggatatgaga gaacttcttc
ccctaagcct cgattcaaga gctatgccta cacacaggct gcttatgtca ccacctctga
ccctacacgg agcccatttc cttcacagca tttggaagct cctgaagaca agtcatttgg
cagttcattg atggagagtg aagtaaacct ggaccgttat caaacagctt tagaagaagt
attatcgtgg cttctttctg ctgaggacac attgcaagca caaggagaga tttctaatga
tgtggaagtg gtgaaagacc agtttcatac tcatgagggg tacatgatgg atttgacagc
ccatcagggc cgggttggta atattctaca attgggaagt aagctgattg gaacaggaaa
attatcagaa gatgaagaaa ctgaagtaca agagcagatg aatctcctaa attcaagatg
ggaatgcctc agggtagcta gcatggaaaa acaaagcaat ttacatagag ttttaatgga
tctccagaat cagaaactga aagagttgaa tgactggcta acaaaaacag aagaaagaac
aaggaaaatg gaggaagagc ctcttggacc tgatcttgaa gacctaaaac gccaagtaca
acaacataag gtgcttcaag aagatctaga acaagaacaa gtcagggtca attctctcac
tcacatggtg gtggtagttg atgaatctag tggagatcac gcaactgctg ctttggaaga
acaacttaag gtattgggag atcgatgggc aaacatctgt agatggacag aagaccgctg
ggttctttta caagacatcc ttctcaaatg gcaacgtctt actgaagaac agtgcctttt
tagtgcatgg ctttcagaaa aagaagatgc agtgaacaag attcacacaa ctggctttaa
agatcaaaat gaaatgttat caagtcttca aaaactggcc gttttaaaag cggatctaga
aaagaaaaag caatccatgg gcaaactgta ttcactcaaa caagatcttc tttcacacct
gaagaataag tcagtgaccc agaagacgga agcatggctg gataactttg cccggtgttg
ggataattta gtccaaaaac ttgaaaagag tacagcacag atttcacag
cag cctgacctag ctcctggact
gaccactatt ggagcctctc ctactcagac tgttactctg gtgacacaac ctgtggttac
taaggaaact gccatctcca aactagaaat gccatcttcc ttgatgttgg aggtacctgc
tctggcagat ttcaaccggg cttggacaga acttaccgac tggctttctc tgcttgatca
agttataaaa tcacagaggg tgatggtggg tgaccttgag gatatcaacg agatgatcat
caagcagaag gcaacaatgc aggatttgga acagaggcgt ccccagttgg aagaactcat
taccgctgcc caaaatttga aaaacaagac cagcaatcaa gaggctagaa caatcattac
ggatcgaatt gaaagaattc agaatcagtg ggatgaagta caagaacacc ttcagaaccg
gaggcaacag ttgaatgaaa tgttaaagga ttcaacacaa tggctggaag ctaaggaaga
agctgagcag gtcttaggac aggccagagc caagcttgag tcatggaagg agggtcccta
tacagtagat gcaatccaaa agaaaatcac agaaaccaag cagttggcca aagacctccg
ccagtggcag acaaatgtag atgtggcaaa tgacttggcc ctgaaacttc tccgggatta
ttctgcagat gataccagaa aagtccacat gataacagag aatatcaatg cctcttggag
aagcattcat aaaagggtga gtgagcgaga ggctgctttg gaagaaactc atagattact
gcaacagttc cccctggacc tggaaaagtt tcttgcctgg cttacagaag ctgaaacaac
tgccaatgtc ctacaggatg ctacccgtaa ggaaaggctc ctagaagact ccaagggagt
aaaagagctg atgaaacaat ggcaagacct ccaaggtgaa attgaagctc acacagatgt
ttatcacaac ctggatgaaa acagccaaaa aatcctgaga tccctggaag gttccgatga
tgcagtcctg ttacaaagac gtttggataa catgaacttc aagtggagtg aacttcggaa
aaagtctctc aacattaggt cccatttgga agccagttct gaccagtgga agcgtctgca
cctttctctg caggaacttc tggtgtggct acagctgaaa gatgatgaat taagccggca
ggcacctatt ggaggcgact ttccagcagt tcagaagcag aacgatgtac atagggcctt
caagagggaa ttgaaaacta aagaacctgt aatcatgagt actcttgaga ctgtacgaat
atttctgaca gagcagcctt tggaaggact agagaaactc taccaggagc ccagagagct

FIGURE 15 (cont.)

gcctcctgag gagagagccc agaatgtcac tcggcttcta cgaaagcagg ctgaggaggt
caatactgag tgggaaaaat tgaacctgca ctccgctgac tggcagagaa aaatagatga
gacccttgaa agactccagg aacttcaaga ggccacggat gagctggacc tcaagctgcg
ccaagctgag gtgatcaagg gatcctggca gcccgtgggc gatctcctca ttgactctct
ccaagatcac ctcgagaaag tcaaggcact tcgaggagaa attgcgcctc tgaaagagaa
cgtgagccac gtcaatgacc ttgctcgcca gcttaccact ttgggcattc agctctcacc
gtataacctc agcactctgg aagacctgaa caccagatgg aagcttctgc aggtggccgt
cgaggaccga gtcaggcagc tgcatgaagc ccacagggac tttggtccag catctcagca
ctttctttcc acgtctgtcc agggtccctg ggagagagcc atctcgccaa acaaagtgcc
ctactatatc aaccacgaga ctcaaacaac ttgctgggac catcccaaaa tgacagagct
ctaccagtct ttagctgacc tgaataatgt cagattctca gcttatagga ctgccatgaa
actccgaaga ctgcagaagg ccctttgctt ggatctcttg agcctgtcag ctgcatgtga
tgccttggac cagcacaacc tcaagcaaaa tgaccagccc atggatatcc tgcagattat
taattgtttg accactattt atgaccgcct ggagcaagag cacaacaatt tggtcaacgt
ccctctctgc gtggatatgt gtctgaactg gctgctgaat gtttatgata cgggacgaac
agggaggatc cgtgtcctgt cttttaaaac tggcatcatt tccctgtgta aagcacattt
ggaagacaag tacagatacc ttttcaagca agtggcaagt tcaacaggat tttgtgacca
gcgcaggctg ggcctccttc tgcatgattc tatccaaatt ccaagacagt tgggtgaagt
tgcatccttt gggggcagta acattgagcc aagtgtccgg agctgcttcc aatttgctaa
taataagcca gagatcgaag cggccctctt cctagactgg atgagactgg aaccccagtc
catggtgtgg ctgcccgtcc tgcacagagt ggctgctgca gaaactgcca agcatcaggc
caaatgtaac atctgcaaag agtgtccaat cattggattc aggtacagga gtctaaagca
ctttaattat gacatctgcc aaagctgctt tttttctggt cgagttgcaa aaggccataa
aatgcactat cccatggtgg aatattgcac tccgactaca tcaggagaag atgttcgaga
ctttgccaag gtactaaaaa acaaatttcg aaccaaaagg tattttgcga agcatccccg
aatgggctac ctgccagtgc agactgtctt agagggggac aacatggaaa ctcccgttac
tctgatcaac ttctggccag tagattctgc gcctgcctcg tcccctcagc tttcacacga
tgatactcat tcacgcattg aacattatgc tagcaggcta gcagaaatgg aaaacagcaa
tggatcttat ctaaatgata gcatctctcc taatgagagc atagatgatg aacatttgtt
aatccagcat tactgccaaa gtttgaacca ggactccccc ctgagccagc ctcgtagtcc
tgcccagatc ttgatttcct tagagagtga ggaaagaggg gagctagaga gaatcctagc
agatcttgag gaagaaaaca ggaatctgca agcagaatat gaccgtctaa agcagcagca
cgaacataaa ggcctgtccc cactgccgtc ccctcctgaa atgatgccca cctctcccca
gagtccccgg gatgctgagc tcattgctga ggccaagcta ctgcgtcaac acaaaggccg
cctggaagcc aggatgcaaa tcctggaaga ccacaataaa cagctggagt cacagttaca
caggctaagg cagctgctgg agcaacccca ggcagaggcc aaagtgaatg gcacaacggt
gtcctctcct tctacctctc tacagaggtc cgacagcagt cagcctatgc tgctccgagt
ggttggcagt caaacttcgg actccatggg tgaggaagat cttctcagtc ctccccagga
cacaagcaca gggttagagg aggtgatgga gcaactcaac aactccttcc ctagttcaag
aggaagaaat acccctggaa agccaatgag agaggacaca atgtaggaag tcttttccac
atggcagatg atttgggcag agcgatggag tccttagtat cagtcatgac agatgaagaa
ggagcagaat aaatgtttta caactcctga ttcccgcatg gtttttataa tattcataca
acaaagagga ttagacagta agagtttaca agaaataaat ctatatttct gtgaagggta
gtggtattat actgtagatt tcagtagttt ctaagtctgt tattgttttg ttaacaatgg
caggttttac acgtctatgc aattgtacaa aaaagttata agaaaactac atgtaaaatc
ttgatagcta aataacttgc catttcttta tatggaacgc attttgggtt gtttaaaaat
ttataacagt tataaagaaa gattgtaaac taaagtgtgc tttataaaaa aaagttgttt
ataaaaaccc ctaaaaacaa aacaaacaca cacacacaca catacacaca cacacacaaa
actttgaggc agcgcattgt tttgcatcct tttggcgtga tatccatatg aaattcatgg
ctttttcttt ttttgcatat taaagataag acttcctcta ccaccacacc aaatgactac
tacacactgc tcatttgaga actgtcagct gagtggggca ggcttgagtt ttcatttcat
atatctatat gtctataagt atataaatac tatagttata tagataaaga gatacgaatt
tctatagact gacttttcc attttttaaa tgttcatgtc acatcctaat agaaagaaat
tacttctagt cagtcatcca ggcttacctg cttggtctag aatggatttt tcccggagcc
ggaagccagg aggaaactac accacactaa aacattgtct acagctccag atgtttctca
ttttaaacaa ctttccactg acaacgaaag taaagtaaag tattggattt ttttaaaggg
aacatgtgaa tgaatacaca ggacttatta tatcagagtg agtaatcggt tggttggttg
attgattgat tgattgatac attcagcttc ctgctgctag caatgccacg atttagattt
aatgatgctt cagtggaaat caatcagaag gtattctgac cttgtgaaca tcagaaggta
ttttttaact cccaagcagt agcaggacga tgatagggct ggagggctat ggattcccag
cccatccctg tgaaggagta ggccactctt taagtgaagg attggatgat tgttcataat
acataaagtt ctctgtaatt acaactaaat tattatgccc tcttctcaca gtcaaaagga

FIGURE 15 (cont.)

actgggtggt ttggtttttg ttgctttttt agatttattg tcccatgtgg gatgagtttt
taaatgccac aagacataat ttaaaataaa taaactttgg gaaaaggtgt aagacagtag
ccccatcaca tttgtgatac tgacaggtat caacccagaa gcccatgaac tgtgtttcca
tcctttgcat ttctctgcga gtagttccac acaggtttgt aagtaagtaa gaaagaaggc
aaattgattc aaatgttaca aaaaaaccct tcttggtgga ttagacaggt taaatatata
aacaaacaaa caaaaattgc tcaaaaaaga ggagaaaagc tcaagaggaa aagctaagga
ctggtaggaa aaagctttac tctttcatgc cattttattt ctttttgatt tttaaatcat
tcattcaata gataccaccg tgtgacctat aattttgcaa atctgttacc tctgacatca
agtgtaatta gcttttggag agtgggctga catcaagtgt aattagcttt tggagagtgg
gttttgtcca ttattaataa ttaattaatt aacatcaaac acggcttctc atgctatttc
tacctcactt tggttttggg gtgttcctga taattgtgca cacctgagtt cacagcttca
ccacttgtcc attgcgttat tttctttttc ctttataatt ctttctttt ccttcataat
tttcaaaaga aaacccaaag ctctaaggta acaaattacc aaattacatg aagatttggt
ttttgtcttg catttttttc ctttatgtga cgctggacct tttctttacc caaggatttt
taaaactcag atttaaaaca aggggttact ttacatccta ctaagaagtt taagtaagta
agtttcattc taaaatcaga ggtaaataga gtgcataaat aattttgttt taatcttttt
gtttttcttt tagacacatt agctctggag tgagtctgtc ataatatttg aacaaaaatt
gagagcttta ttgctgcatt ttaagcataa ttaatttgga cattatttcg tgttgtgttc
tttataacca ccgagtatta aactgtaaat cataatgtaa ctgaagcata aacatcacat
ggcatgtttt gtcattgttt tcaggtactg agttcttact tgagtatcat aatatattgt
gttttaacac caacactgta acatttacga attatttttt taaacttcag ttttactgca
ttttcacaac atatcagact tcaccaaata tatgccttac tattgtatta tagtactgct
ttactgtgta tctcaataaa gcacgcagtt atgttac

FIGURE 16 (SEQ ID NO:87)

Human skeletal muscle alpha actinin, complete cDNA sequence:

Genbank accession # M86406; 4181 base pairs

GGAACTCCGCTTCGCCCGAGACCCAGCGCCCAGGCGTGTCGCCCGAGAGGAGCCGCGCGAAG
GTCACCCCGCGCCCGCCGCCCGCCGCCCGCCGCCTCCGTGGGTCCGTTTGCCAGTCAGCCCGT
GCGTCCGAGCCCCTCGCGCCCCGCCGCAGCCCCGGCCAACCGAGCGCCATGAACCAGATAGA
GCCCGGCGTGCAGTACAACTACGTGTACGACGAGGATGAGTACATGATCCAGGAGGAGGAGT
GGGACCGCGACCTGCTCCTGGACCCAGCCTGGGAGAAGCAGCAGAGGAAGACCTTCACTGCC
TGGTGTAACTCCCACCTAAGGAAAGCCGGCACCCAGATTGAGAACATCGAGGAAGACTTCAG
GAATGGCCTTAAGCTCATGCTGCTTTTGGAAGTCATCTCAGGGGAAAGGCTGCCCAAACCTGA
CCGGGGAAAAATGCGGTTCCACAAAATTGCTAATGTCAACAAAGCTTTGGATTACATAGCCA
GCAAAGGGGTGAAACTGGTGTCCATCGGCGCTGAAGAAATTGTTGATGGCAATGTGAAAATG
ACCCTGGGTATGATCTGGACCATCATCCTTCGCTTTGCTATTCAGGATATTTCGGTTGAAGAAA
CATCTGCCAAAGAAGGTCTGCTGCTTTGGTGTCAGAGGAAAACTGCTCCTTATAGAAATGTGA
ACATTCAGAACTTCCATACTAGCTGGAAAGATGGCCTTGGACTCTGTGCCCTCATCCACCGAC
ACCGGCCTGACCTCATTGACTACTCAAAGCTTAACAAGGATGACCCCATAGGAAATATTAACC
TGGCCATGGAAATCGCTGAGAAGCACCTGGATATTCCTAAAATGTTGGATGCTGAAGACATCG
TGAACACCCCTAAACCCGATGAAAGAGCCATCATGACGTACGTCTCTTGCTTCTACCACGCTT
TTGCGGGCGCGGAGCAGGCCGAGACAGCGGCTAACAGGATATGTAAGGTTCTTGCTGTGAAT
CAAGAGAATGAGAGGCTGATGGAAGAATATGAGAGGCTAGCGAGTGAGCTTTTGGAATGGAT
TCGTCGCACGATCCCCTGGCTGGAGAACCGGACTCCCGAGAAGACCATGCAAGCCATGCAGA
AGAAGCTGGAGGACTTCCGGGATTACCGCCGGAAGCACAAGCCACCCAAGGTGCAGGAGAA
ATGCCAGCTGGAGATCAACTTCAACACGCTGCAGACCAAGCTGCGGATCAGCAACCGTCCTG
CCTTCATGCCCTCCGAGGGCAAGATGGTGTCGGATATTGCTGGTGCCTGGCAGAGGCTGGAGC
AGGCTGAGAAGGGTTACGAGGAGTGGTTGCTCAATGAGATTCGGAGACTGGAGCGCTTGGAA
CACCTGGCTGAGAAGTTCAGGCAGAAGGCCTCAACGCACGAGACTTGGGCTTATGGCAAAGA
GCAGATCTTGCTGCAGAAGGATTACGAGTCGGCGTCGCTGACAGAGGTGCGGGCTCTGCTGC
GGAAGCACGAGGCGTTCGAGAGCGACCTGGCAGCGCACCAGGACCGCGTGGAGCAGATCGC
AGCCATCGCGCAGGAGCTCAATGAACTGGACTATCACGACGCTGTGAATGTCAATGATCGGT
GCCAGAAAATTTGTGACCAGTGGGACCGACTGGGAACGCTTACTCAGAAGAGGAGAGAAGCC
CTAGAGAGAATGGAGAAATTGCTAGAAACCATTGATCAGCTTCACCTGGAGTTTGCCAAGAG
GGCTGCTCCTTTCAACAATTGGATGGAGGGCGCTATGGAGGATCTGCAAGATATGTTCATTGT
CCACAGCATTGAGGAGATCCAGAGTCTGATCACTGCGCATGAGCAGTTCAAGGCCACGCTGC
CCGAGGCGGACGGAGAGCGGCAGTCCATCATGGCCATCCAGAACGAGGTGGAGAAGGTGATT
CAGAGCTACAACATCAGAATCAGCTCAAGCAACCCGTACAGCACTGTCACCATGGATGAGCT
CCGGACCAAGTGGGACAAGGTGAAGCAACTCGTGCCCATCCGCGATCAATCCCTGCAGGAGG
AGCTGGCTCGCCAGCATGCTAACGAGCGTCTGAGGCGCCAGTTTGCTGCCCAAGCCAATGCCA
TTGGGCCCTGGATCCAGAACAAGATGGAGGAGATTGCCCGGAGCTCCATCCAGATCACAGGA
GCCCTGGAAGACCAGATGAACCAGCTGAAGCAGTATGAGCACAACATCATCAACTATAAGAA
CAACATCGACAAGCTGGAGGGAGACCATCAGCTCATCCAGGAGGCCCTTGTCTTTGACAACA
AGCACACGAACTACACGATGGAGCACATTCGTGTTGGATGGGAGCTGCTGCTGACAACCATC
GCCAGAACCATCAATGAGGTGGAGACTCAGATCCTGACGAGAGATGCGAAGGGCATCACCCA
GGAGCAGATGAATGAGTTCAGAGCCTCCTTCAACCACTTTGACAGGAGGAAGAATGGCCTGA
TGGATCATGAGGATTTCAGAGCCTGCCTGATTTCCATGGGTTATGACCTGGGTGAAGCCGAAT
TTGCCCGCATTATGACCCTGGTAGATCCCAACGGGCAAGGCACCGTCACCTTCCAATCCTTCA
TCGACTTCATGACTAGAGAGACGGCTGACACCGACACTGCCGAGCAGGTCATCGCCTCCTTCC
GGATCCTGGCTTCTGATAAGCCATACATCCTGGCGGAGGAGCTGCGTCGGGAGCTGCCCCCGG
ATCAGGCCCAGTACTGCATCAAGAGGATGCCCGCCTACTCGGGCCCAGGCAGTGTGCCTGGTG
CACTGGATTACGCTGCGTTCTCTTCCGCACTCTACGGGGAGAGCGATCTGTGATGCTGAGCTT
CTGTAATCACTCATCCCATCAGAATGCAATAAAAGCGGAAGTCACAGTTTGTTTCCTGGAAAC
TTTGACAAGCTTTATTAAGTTGAGAGAGAGAGAGGGGGAAAAAAAAAAAAGCCTTTCGTAGTT
CAGTAATTGCCAGCAATATAACACGGCTAAAATGAAGTTTTTACAGTATATGACATAGTGCGC

FIGURE 16 (cont.)

TTCATAAATAGGTTTATTTCTGAGTTTTTAGCAAAATGTAATGAAATATCAGGTTGATTTCTTT
GATTAAACAGAACAAATTACTTGAGTAATAGGAAATTAGGAGGATCTAGGGACAGAAGGAAA
GTGAAAAATGTGAAAATACAAAATACCCAAGATTTAAGACCGGGGGGAAAAAACCACAAATT
GGTAAATAAAGGTTTGCTATTTGTAAAAAATTTCATTTATCTCTAATATGCTTATGTGATTGGC
CCTAGGGGAGTATATTTGGGATTCTAATGTTTTATTTTCATGCTTATCCAAAGATTACTATTGT
ATCTTCAAATGAACTTAATATTGTGAGATGGAACTGCCGGGGATTAAAAAGACTACCCAAAA
GATTTTTGGCACTTACAATTTTTAAAATAGTTTATGTCATCTCTTCATTATTTAGGGCTGGATG
GTCAACTCAGTCAGTGATTTTTTGATGCTTCTCTTATCCTCCAGAATAGAGACCTAAGGACACG
TGGAAGTCAGTTTAATTGCCAGAGAGAAGGATGCAATCACTAGGTGAAATGAGGTTTTTAGG
ATTATTTATTGATTCCAGGTTCCCATGCTTTTTGTTAGAGCTTATTAGTACAGGTTCTCAAGAG
ATGACCACATAAAAGTGCTCTGTTTATAAATAAGCAGGTTTCTGTAGTACTGACTGGTTCATC
ACAAGGCAAGTCAGAAACCAGTATCCTTCTAGCTCTCCAGTCAGGACTTCCTTATGCCTCTAG
TTTTATGACCGGTTAAGGAGAAGCCAGAGTTAGAGTAGGAGAGGACTAATTCTCAGCAGCAG
TGGAGGTGAGTTCTTTCTTTTGCGGAAGCTTTACATATGTTTTGTGTAGTAGGAATAACTAGAT
ATTTTAGCTAGTGTGCGGTGTGTGTTCACCCCTGGGATTGGACAGTGTATCCTAACAAGTCCC
ATGTCTGGTTCTGTGTCTAAAGGCCTGCTCCATGACACAGGATGCTACATGCACTCCTGCTAG
CACATCTTGATCTGTTGAATGTTCATTCTTTCTTTTTGCTCATACTGCTGTAGGCTATAATTCCC
CCCTGTTTTTCCATCTTGTTGACAGCTTGTAGAGAATAAAGCAGGAATTC

FIGURE 17 (16-repeat construct, SEQ ID NO:44) (numbering corresponds to the numbering of human dystrophin, acc. no. M185330

```
   1 gggattccct cactttcccc ctacaggact cagatctggg aggcaattac cttcggagaa
  61 aaacgaatag gaaaaactga agtgttactt tttttaaagc tgctgaagtt tgttggtttc
 121 tcattgtttt taagcctact ggagcaataa agtttgaaga acttttacca ggtttttttt
 181 atcgctgcct tgatatacac ttttcaaaat gctttggtgg gaagaagtag aggactgtta
 241 tgaaagagaa gatgttcaaa agaaaacatt cacaaaatgg gtaaatgcac aattttctaa
 301 gtttgggaag cagcatattg agaacctctt cagtgaccta caggatggga ggcgcctcct
 361 agacctcctc gaaggcctga cagggcaaaa actgccaaaa gaaaaaggat ccacaagagt
 421 tcatgccctg aacaatgtca acaaggcact gcgggttttg cagaacaata atgttgattt
 481 agtgaatatt ggaagtactg acatcgtaga tggaaatcat aaactgactc ttggtttgat
 541 ttggaatata atcctccact ggcaggtcaa aaatgtaatg aaaaatatca tggctggatt
 601 gcaacaaacc aacagtgaaa agattctcct gagctgggtc cgacaatcaa ctcgtaatta
 661 tccacaggtt aatgtaatca acttcaccac cagctggtct gatggcctgg ctttgaatgc
 721 tctcatccat agtcataggc cagacctatt tgactggaat agtgtggttt gccagcagtc
 781 agccacacaa cgactggaac atgcattcaa catcgccaga tatcaattag gcatagagaa
 841 actactcgat cctgaagatg ttgataccac ctatccagat aagaagtcca tcttaatgta
 901 catcacatca ctcttccaag ttttgcctca acaagtgagc attgaagcca tccaggaagt
 961 ggaaatgttg ccaaggccac ctaaagtgac taaagaagaa cattttcagt tacatcatca
1021 aatgcactat tctcaacaga tcacggtcag tctagcacag ggatatgaga gaacttcttc
1081 ccctaagcct cgattcaaga gctatgccta cacacaggct gcttatgtca ccacctctga
1141 ccctacacgg agcccatttc cttcacagca tttggaagct cctgaagaca agtcatttgg
1201 cagttcattg atggagagtg aagtaaacct ggaccgttat caaacagctt tagaagaagt
1261 attatcgtgg cttctttctg ctgaggacac attgcaagca caaggagaga tttctaatga
1321 tgtggaagtg gtgaaagacc agtttcatac tcatgagggg tacatgatgg atttgacagc
1381 ccatcagggc cgggttggta atattctaca attgggaagt aagctgattg gaacaggaaa
1441 attatcagaa gatgaagaaa ctgaagtaca agagcagatg aatctcctaa attcaagatg
1501 ggaatgcctc agggtagcta gcatggaaaa acaaagcaat ttacatagag ttttaatgga
1561 tctccagaat cagaaactga aagagttgaa tgactggcta acaaaaacag aagaaagaac
1621 aaggaaaatg gaggaagagc ctcttggacc tgatcttgaa gacctaaaac gccaagtaca
1681 acaacataag gtgcttcaag aagatctaga acaagaacaa gtcagggtca attctctcac
1741 tcacatggtg gtggtagttg atgaatctag tggagatcac gcaactgctg ctttggaaga
1801 acaacttaag gtattgggag atcgatgggc aaacatctgt agatggacag aagaccgctg
1861 ggttctttta caagacatcc ttctcaaatg gcaacgtctt actgaagaac agtgcctttt
1921 tagtgcatgg ctttcagaaa aagaagatgc agtgaacaag attcacacaa ctggctttaa
1981 agatcaaaat gaaatgttat caagtcttca aaaactggcc gttttaaaag cggatctaga
2041 aaagaaaaag caatccatgg gcaaactgta ttcactcaaa caagatcttc tttcaacact
2101 gaagaataag tcagtgaccc agaagacgga agcatggctg gataactttg cccggtgttg
2161 ggataattta gtccaaaaac ttgaaaagag tacagcacag atttcacagg ctgtcaccac
2221 cactcagcca tcactaacac agacaactgt aatggaaaca gtaactacgg tgaccacaag
2281 ggaacagatc ctggtaaagc atgctcaaga ggaacttcca ccaccacctc cccaaaagaa
2341 gaggcagatt actgtggatt ctgaaattag gaaaaggttg gatgttgata taactgaact
2401 tcacagctgg attactcgct cagaagctgt gttgcagagt cctgaatttg caatctttcg
2461 gaaggaaggc aacttctcag acttaaaaga aaaagtcaat gccatagagc gagaaaaagc
2521 tgagaagttc agaaaactgc aagatgccag cagatcagct caggccctgg tggaacagat
2581 ggtgaatgag ggtgttaatg cagatagcat caaacaagcc tcagaacaac tgaacagccg
2641 gtggatcgaa ttctgccagt tgctaagtga gagacttaac tggctggagt atcagaacaa
2701 catcatcgct ttctataatc agctacaaca attggagcag atgacaacta ctgctgaaaa
2761 ctggttgaaa atccaaccca ccaccccatc agagccaaca gcaattaaaa gtcagttaaa
2821 aatttgtaag gatgaagtca accggctatc aggtcttcaa cctcaaattg aacgattaaa
2881 aattcaaagc atagccctga aagagaaagg acaaggaccc atgttcctgg atgcagactt
2941 tgtggccttt acaaatcatt ttaagcaagt cttttctgat gtgcaggcca gagagaaaga
3001 gctacagaca atttttgaca ctttgccacc aatgcgctat caggagacca tgagtgccat
3061 caggacatgg gtccagcagt cagaaaccaa actctccata cctcaactta gtgtcaccga
3121 ctatgaaatc atggagcaga gactcgggga attgcaggct ttacaaagtt ctctgcaaga
3181 gcaacaaagt ggcctatact atctcagcac cactgtgaaa gagatgtcga agaaagcgcc
3241 ctctgaaatt agccggaaat atcaatcaga atttgaagaa attgagggac gctggaagaa
3301 gctctcctcc cagctggttg agcattgtca aaagctagag gagcaaatga ataaactccg
3361 aaaaattcag aatcacatac aaaccctgaa gaaatggatg gctgaagttg atgtttttct
3421 gaaggaggaa tggcctgccc ttggggattc agaaattcta aaaaagcagc tgaaacagtg
3481 cagactttta gtcagtgata ttcagacaat tcagcccagt ctaaacagtg tcaatgaagg
3541 tgggcagaag ataaagaatg aagcagagcc agagtttgct tcgagacttg agacagaact
3601 caaagaactt aacactcagt gggatcacat gtgccaacag gtctatgcca gaaaggaggc
```

FIGURE 17 (cont.)

```
3661 cttgaaggga ggtttggaga aaactgtaag cctccagaaa gatctatcag agatgcacga
3721 atggatgaca caagctgaag aagagtatct tgagagagat tttgaatata aaactccaga
3781 tgaattacag aaagcagttg aagagatgaa gagagctaaa gaagaggccc aacaaaaaga
3841 agcgaaagtg aaactcctta ctgagtctgt aaatagtgtc atagctcaag ctccacctgt
3901 agcacaagag gccttaaaaa aggaacttga aactctaacc accaactacc agtggctctg
3961 cactaggctg aatgggaaat gcaagacttt ggaagaa
6512 tctgttgag aaatggcggc gttttcatta
6541 tgatataaag atatttaatc agtggctaac agaagctgaa cagtttctca gaaagacaca
6601 aattcctgag aattgggaac atgctaaata caaatggtat cttaaggaac tccaggatgg
6661 cattgggcag cggcaaactg ttgtcagaac attgaatgca actggggaag aaataattca
6721 gcaatcctca aaaacagatg ccagtattct acaggaaaaa ttgggaagcc tgaatctgcg
6781 gtggcaggag gtctgcaaac agctgtcaga cagaaaaaag aggctagaag aacaaaagaa
6841 tatcttgtca gaatttcaaa gagatttaaa tgaatttgtt ttatggttgg aggaagcaga
6901 taacattgct agtatcccac ttgaacctgg aaaagagcag caactaaaag aaaagcttga
6961 gcaagtcaag ttactggtgg aagagttgcc cctgcgccag ggaattctca aacaattaaa
7021 tgaaactgga ggacccgtgc ttgtaagtgc tcccataagc ccagaagagc aagataaact
7081 tgaaaataag ctcaagcaga caaatctcca gtggataaag gtttccagag ctttacctga
7141 gaaacaagga gaaattgaag ctcaaataaa agaccttggg cagcttgaaa aaaagcttga
7201 agaccttgaa gagcagttaa atcatctgct gctgtggtta tctcctatta ggaatcagtt
7261 ggaaatttat aaccaaccaa accaagaagg accatttgac gttcaggaaa ctgaaatagc
7321 agttcaagct aaacaaccgg atgtggaaga gattttgtct aaagggcagc atttgtacaa
7381 ggaaaaacca gccactcagc cagtgaagag gaagttagaa gatctgagct ctgagtggaa
7441 ggcggtaaac cgtttacttc aagagctgag ggcaaagcag cctgacctag ctcctggact
7501 gaccactatt ggagcctctc ctactcagac tgttactctg gtgacacaac ctgtggttac
7561 taaggaaact gccatctcca aactagaaat gccatcttcc ttgatgttgg aggtacctgc
7621 tctggcagat ttcaaccggg cttggacaga acttaccgac tggctttctc tgcttgatca
7681 agttataaaa tcacagaggg tgatggtggg tgaccttgag gatatcaacg agatgatcat
7741 caagcagaag gcaacaatgc aggatttgga acagaggcgt ccccagttgg aagaactcat
7801 taccgctgcc caaaatttga aaaacaagac cagcaatcaa gaggctagaa caatcattac
7861 ggatcgaatt gaaagaattc agaatcagtg ggatgaagta caagaacacc ttcagaaccg
7921 gaggcaacag ttgaatgaaa tgttaaagga ttcaacacaa tggctggaag ctaaggaaga
7981 agctgagcag gtcttaggac aggccagagc caagcttgag tcatggaagg agggtcccta
8041 tacagtagat gcaatccaaa agaaaatcac agaaaccaag cagttggcca aagacctccg
8101 ccagtggcag acaaatgtag atgtggcaaa tgacttggcc ctgaaacttc tccgggatta
8161 ttctgcagat gataccagaa aagtccacat gataacagag aatatcaatg cctcttggag
8221 aagcattcat aaaagggtga gtgagcgaga ggctgctttg gaagaaactc atagattact
8281 gcaacagttc cccctggacc tggaaaagtt tcttgcctgg cttacagaag ctgaaacaac
8341 tgccaatgtc ctacaggatg ctacccgtaa ggaaaggctc ctagaagact ccaagggagt
8401 aaaagagctg atgaaacaat ggcaagacct ccaaggtgaa attgaagctc acacagatgt
8461 ttatcacaac ctggatgaaa acagccaaaa aatcctgaga tccctggaag gttccgatga
8521 tgcagtcctg ttacaaagac gtttggataa catgaacttc aagtggagtg aacttcggaa
8581 aaagtctctc aacattaggt cccatttgga agccagttct gaccagtgga agcgtctgca
8641 cctttctctg caggaacttc tggtgtggct acagctgaaa gatgatgaat taagccggca
8701 ggcacctatt ggaggcgact ttccagcagt tcagaagcag aacgatgtac atagggcctt
8761 caagagggaa ttgaaaacta aagaacctgt aatcatgagt actcttgaga ctgtacgaat
8821 atttctgaca gagcagcctt tggaaggact agagaaactc taccaggagc ccagagagct
8881 gcctcctgag gagagagccc agaatgtcac tcggcttcta cgaaagcagg ctgaggaggt
8941 caatactgag tgggaaaaat tgaacctgca ctccgctgac tggcagagaa aaatagatga
9001 gacccttgaa agactccagg aacttcaaga ggccacggat gagctggacc tcaagctgcg
9061 ccaagctgag gtgatcaagg gatcctggca gcccgtgggc gatctcctca ttgactctct
9121 ccaagatcac ctcgagaaag tcaaggcact tcgaggagaa attgcgcctc tgaaagagaa
9181 cgtgagccac gtcaatgacc ttgctcgcca gcttaccact ttgggcattc agctctcacc
9241 gtataacctc agcactctgg aagacctgaa caccagatgg aagcttctgc aggtggccgt
9301 cgaggaccga gtcaggcagc tgcatgaagc ccacagggac tttggtccag catctcagca
9361 ctttctttcc acgtctgtcc agggtccctg ggagagagcc atctcgccaa acaaagtgcc
9421 ctactatatc aaccacgaga ctcaaacaac ttgctgggac catcccaaaa tgacagagct
9481 ctaccagtct ttagctgacc tgaataatgt cagattctca gcttatagga ctgccatgaa
9541 actccgaaga ctgcagaagg ccctttgctt ggatctcttg agcctgtcag ctgcatgtga
9601 tgccttggac cagcacaacc tcaagcaaaa tgaccagccc atggatatcc tgcagattat
9661 taattgtttg accactattt atgaccgcct ggagcaagag cacaacaatt tggtcaacgt
9721 ccctctctgc gtggatatgt gtctgaactg gctgctgaat gtttatgata cgggacgaac
9781 agggaggatc cgtgtcctgt cttttaaaac tggcatcatt tccctgtgta aagcacattt
9841 ggaagacaag tacagatacc ttttcaagca agtggcaagt tcaacaggat tttgtgacca
9901 gcgcaggctg ggcctccttc tgcatgattc tatccaaatt ccaagacagt tgggtgaagt
9961 tgcatccttt gggggcagta acattgagcc aagtgtccgg agctgcttcc aatttgctaa
10021 taataagcca gagatcgaag cggccctctt cctagactgg atgagactgg aaccccagtc
10081 catggtgtgg ctgcccgtcc tgcacagagt ggctgctgca gaaactgcca agcatcaggc
10141 caaatgtaac atctgcaaag agtgtccaat cattggattc aggtacagga gtctaaagca
```

6023884.1000500

41/66.

FIGU · 17 (cont.)

```
10201 ctttaatcat gacatctgcc aaagctgctt ttttctcygt cgagttgcaa aaggccataa
10261 aatgcactat cccatggtgg aatattgcac tccgactaca tcaggagaag atgttcgaga
10321 ctttgccaag gtactaaaaa acaaatttcg aaccaaaagg tattttgcga agcatccccg
10381 aatgggctac ctgccagtgc agactgtctt agagggggac aacatggaaa ctcccgttac
10441 tctgatcaac ttctggccag tagattctgc gcctgcctcg tcccctcagc tttcacacga
10501 tgatactcat tcacgcattg aacattatgc tagcaggcta gcagaaatgg aaaacagcaa
10561 tggatcttat ctaaatgata gcatctctcc taatgagagc atagatgatg aacatttgtt
10621 aatccagcat tactgccaaa gtttgaacca ggactccccc ctgagccagc ctcgtagtcc
10681 tgcccagatc ttgatttcct tagagagtga ggaaagaggg gagctagaga gaatcctagc
10741 agatcttgag gaagaaaaca ggaatctgca agcagaatat gaccgtctaa agcagcagca
10801 cgaacataaa ggcctgtccc cactgccgtc ccctcctgaa atgatgccca cctctcccca
10861 gagtccccgg gatgctgagc tcattgctga ggccaagcta ctgcgtcaac acaaaggccg
10921 cctggaagcc aggatgcaaa tcctggaaga ccacaataaa cagctggagt cacagttaca
10981 caggctaagg cagctgctgg agcaacccca ggcagaggcc aaagtgaatg gcacaacggt
11041 gtcctctcct tctacctctc tacagaggtc cgacagcagt cagcctatgc tgctccgagt
11101 ggttggcagt caaacttcgg actccatggg tgaggaagat cttctcagtc ctccccagga
11161 cacaagcaca gggttagagg aggtgatgga gcaactcaac aactccttcc ctagttcaag
11221 aggaagaaat acccctggaa agccaatgag agaggacaca atgtaggaag tcttttccac
11281 atggcagatg atttgggcag agcgatggag tccttagtat cagtcatgac agatgaagaa
11341 ggagcagaat aaatgtttta caactcctga ttcccgcatg gtttttataa tattcataca
11401 acaaagagga ttagacagta agagtttaca agaaataaat ctatattttt gtgaagggta
11461 gtggtattat actgtagatt tcagtagttt ctaagtctgt tattgttttg ttaacaatgg
11521 caggttttac acgtctatgc aattgtacaa aaaagttata agaaaactac atgtaaaatc
11581 ttgatagcta aataacttgc catttcttta tatggaacgc attttgggtt gtttaaaaat
11641 ttataacagt tataaagaaa gattgtaaac taaagtgtgc tttataaaaa aaagttgttt
11701 ataaaaaccc ctaaaaacaa aacaaacaca cacacacaca catacacaca cacacacaaa
11761 actttgaggc agcgcattgt tttgcatcct tttggcgtga tatccatatg aaattcatgg
11821 ctttttcttt ttttgcatat taaagataag acttcctcta ccaccacacc aaatgactac
11881 tacacactgc tcatttgaga actgtcagct gagtggggca ggcttgagtt ttcatttcat
11941 atatctatat gtctataagt atataaatac tatagttata tagataaaga gatacgaatt
12001 tctatagact gacttttcc atttttaaa tgttcatgtc acatcctaat agaaagaaat
12061 tacttctagt cagtcatcca ggcttacctg cttggtctag aatggatttt tcccggagcc
12121 ggaagccagg aggaaactac accacactaa aacattgtct acagctccag atgtttctca
12181 ttttaaacaa ctttccactg acaacgaaag taaagtaaag tattggattt ttttaaaggg
12241 aacatgtgaa tgaatacaca ggacttatta tatcagagtg agtaatcggt tggttggttg
12301 attgattgat tgattgatac attcagcttc ctgctgctag caatgccacg atttagattt
12361 aatgatgctt cagtggaaat caatcagaag gtattctgac cttgtgaaca tcagaaggta
12421 ttttttaact cccaagcagt agcaggacga tgatagggct ggagggctat ggattcccag
12481 cccatccctg tgaaggagta ggccactctt taagtgaagg attggatgat tgttcataat
12541 acataaagtt ctctgtaatt acaactaaat tattatgccc tcttctcaca gtcaaaagga
12601 actgggtggt ttggtttttg ttgctttttt agatttattg tcccatgtgg gatgagtttt
12661 taaatgccac aagacataat ttaaaataaa taaactttgg gaaaaggtgt aagacagtag
12721 ccccatcaca tttgtgatac tgacaggtat caacccagaa gcccatgaac tgtgtttcca
12781 tcctttgcat ttctctgcga gtagttccac acaggtttgt aagtaagtaa gaaagaaggc
12841 aaattgattc aaatgttaca aaaaaaccct tcttggtgga ttagacaggt taaatatata
12901 aacaaacaaa caaaaattgc tcaaaaaaga ggagaaaagc tcaagaggaa aagctaagga
12961 ctggtaggaa aaagctttac tctttcatgc cattttattt ctttttgatt tttaaatcat
13021 tcattcaata gataccaccg tgtgacctat aattttgcaa atctgttacc tctgacatca
13081 agtgtaatta gcttttggag agtgggctga catcaagtgt aattagcttt tggagagtgg
13141 gttttgtcca ttattaataa ttaattaatt aacatcaaac acggcttctc atgctatttc
13201 tacctcactt tggttttggg gtgttcctga taattgtgca cacctgagtt cacagcttca
13261 ccacttgtcc attgcgttat tttctttttc ctttataatt ctttcttttt ccttcataat
13321 tttcaaaaga aaacccaaag ctctaaggta acaaattacc aaattacatg aagatttggt
13381 ttttgtcttg catttttttc ctttatgtga cgctggacct tttctttacc caaggatttt
13441 taaaactcag atttaaaaca aggggttact ttacatccta ctaagaagtt taagtaagta
13501 agtttcattc taaaatcaga ggtaaataga gtgcataaat aattttgttt taatcttttt
13561 gtttttcttt tagacacatt agctctggag tgagtctgtc ataatatttg aacaaaaatt
13621 gagagcttta ttgctgcatt ttaagcataa ttaatttgga cattatttcg tgttgtgttc
13681 tttataacca ccgagtatta aactgtaaat cataagtaa ctgaagcata aacatcacat
13741 ggcatgtttt gtcattgttt tcaggtactg agttcttact tgagtatcat aatatattgt
13801 gttttaacac caacactgta acatttacga attatttttt taaacttcag ttttactgca
13861 ttttcacaac atatcagact tcaccaaata tatgccttac tattgtatta tagtactgct
13921 ttactgtgta tctcaataaa gcacgcagtt atgttac
```

FIGURE 18 (WW domain, SEQ ID NO:45)

```
9371 acgtctgtcc agggtccctg ggagagagcc atctcgccaa acaaagtgcc
9421 ctactatatc aaccacgaga ctcaaacaac ttgctgggac catcccaaaa tgacagagct
9481 ctac
```

FIGURE 19

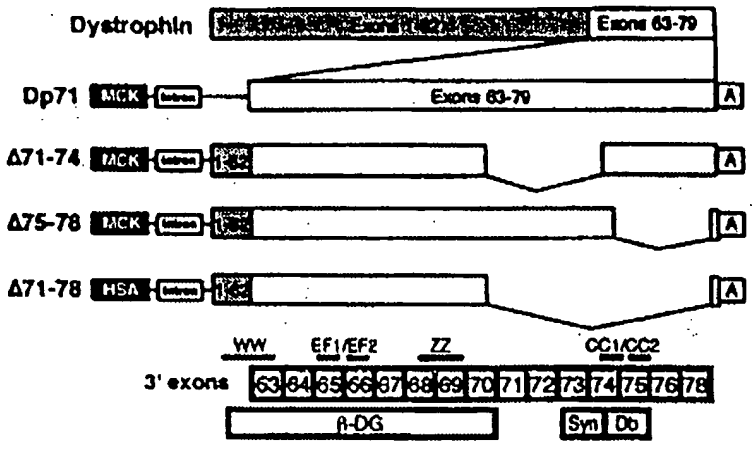
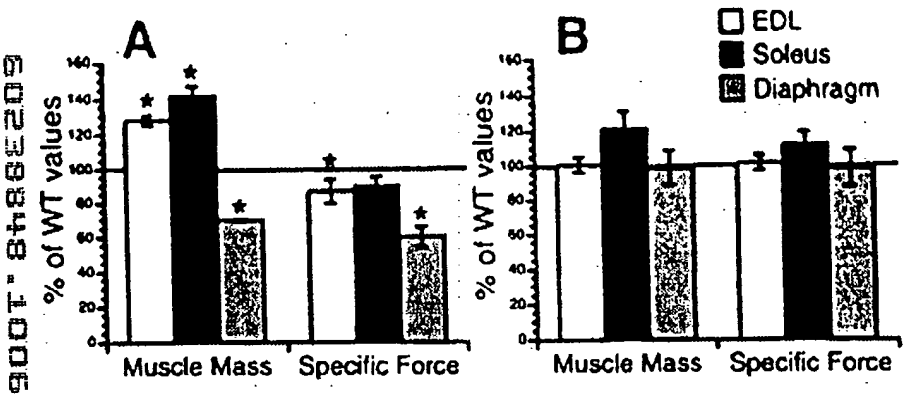


FIGURE 20



Contractile properties of EDL, soleus, and diaphragm muscles in wild-type, *mdx*, and dystrophin Δ71–78 mice. Muscle mass and specific force for mdx (A) and Δ71–78 (B) muscles were charted as a percentage of wild-type values. Significant differences ($P < 0.05$) are marked with an asterisk (*).

FIGURE 21 (pBSX sequence, SEQ ID NO:46)

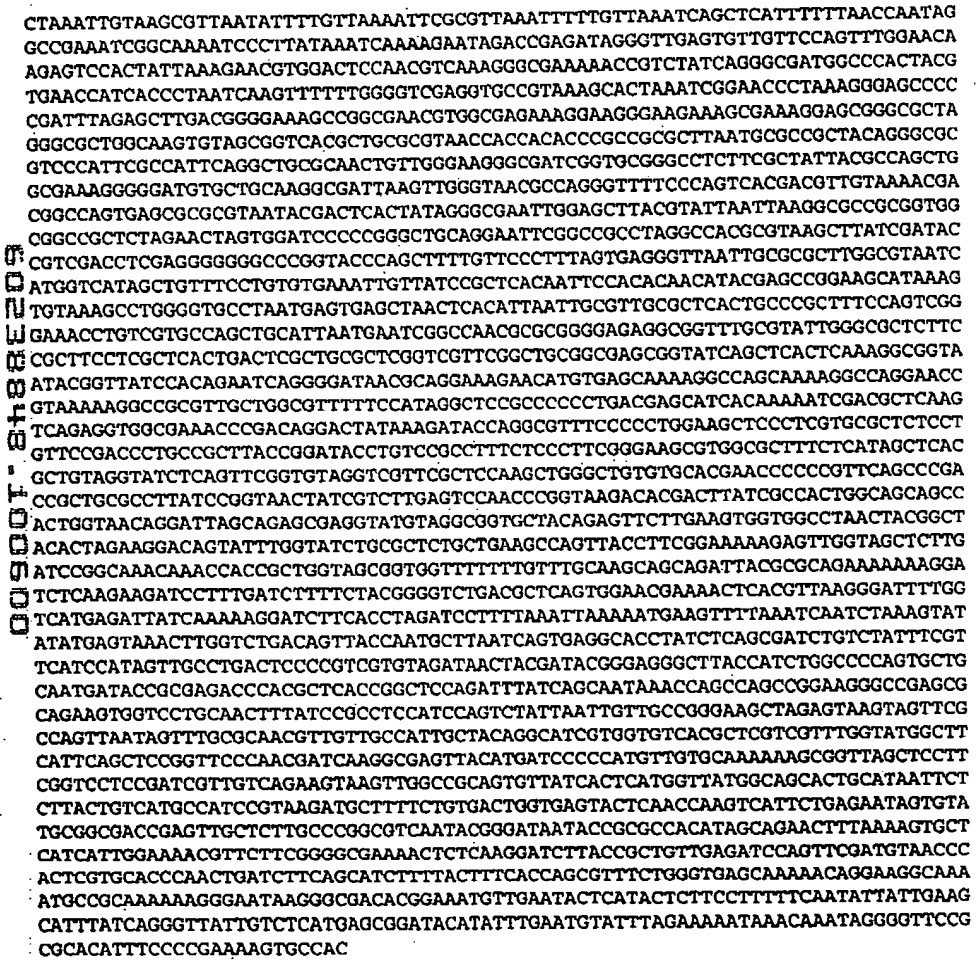


FIGURE 22

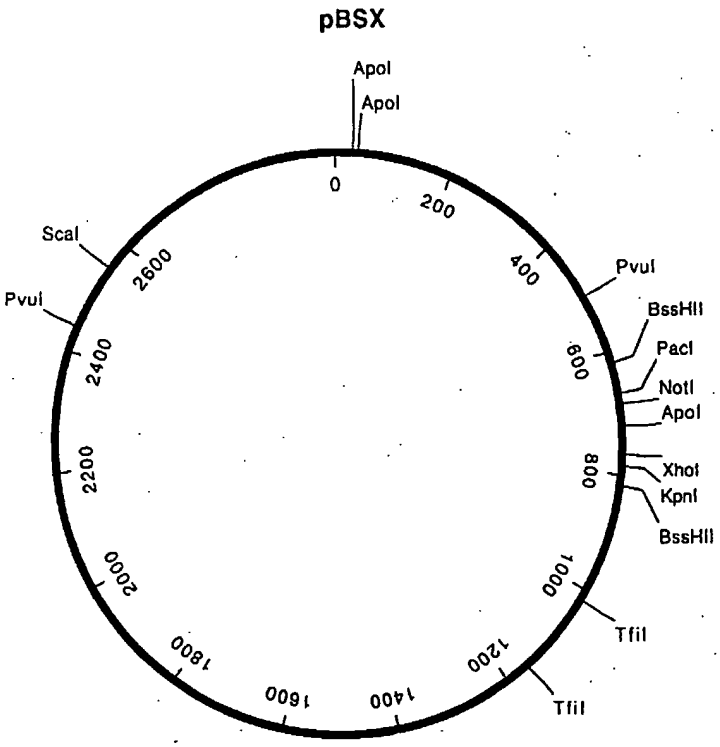

pBSX
ApoI
ApoI
0
200
400
PvuI
BssHII
PacI
NotI
ApoI
600
800
XhoI
KpnI
BssHII
1000
TfiI
1200
TfiI
1400
1600
1800
2000
2200
2400
PvuI
ScaI
2600

FIGURE 23 ("full-length' HDMD, SEQ ID NO:47)
-numbering corresponds to human dystrophin SEQ ID NO:1

```
   1 gggattccct cactttcccc ctacaggact cagatctggg aggcaattac cttcggagaa
  61 aaacgaatag gaaaaactga agtgttactt tttttaaagc tgctgaagtt tgttggtttc
 121 tcattgtttt taagcctact ggagcaataa agtttgaaga acttttacca ggttttttt
 181 atcgctgcct tgatatacac ttttcaaaat gctttggtgg gaagaagtag aggactgtta
 241 tgaaagagaa gatgttcaaa agaaaacatt cacaaaatgg gtaaatgcac aattttctaa
 301 gtttgggaag cagcatattg agaacctctt cagtgaccta caggatggga ggcgcctcct
 361 agacctcctc gaaggcctga cagggcaaaa actgccaaaa gaaaaaggat ccacaagagt
 421 tcatgccctg aacaatgtca acaaggcact gcgggttttg cagaacaata atgttgattt
 481 agtgaatatt ggaagtactg acatcgtaga tggaaatcat aaactgactc ttggtttgat
 541 ttggaatata atcctccact ggcaggtcaa aaatgtaatg aaaaatatca tggctggatt
 601 gcaacaaacc aacagtgaaa agattctcct gagctgggtc cgacaatcaa ctcgtaatta
 661 tccacaggtt aatgtaatca acttcaccac cagctggtct gatggcctgg ctttgaatgc
 721 tctcatccat agtcataggc cagacctatt tgactggaat agtgtggttt gccagcagtc
 781 agccacacaa cgactggaac atgcattcaa catcgccaga tatcaattag gcatagagaa
 841 actactcgat cctgaagatg ttgataccac ctatccagat aagaagtcca tcttaatgta
 901 catcacatca ctcttccaag ttttgcctca acaagtgagc attgaagcca tccaggaagt
 961 ggaaatgttg ccaaggccac ctaaagtgac taaagaagaa cattttcagt tacatcatca
1021 aatgcactat tctcaacaga tcacggtcag tctagcacag ggatatgaga gaacttcttc
1081 ccctaagcct cgattcaaga gctatgccta cacacaggct gcttatgtca ccacctctga
1141 ccctacacgg agcccatttc cttcacagca tttggaagct cctgaagaca agtcatttgg
1201 cagttcattg atggagagtg aagtaaacct ggaccgttat caaacagctt tagaagaagt
1261 attatcgtgg cttctttctg ctgaggacac attgcaagca caaggagaga tttctaatga
1321 tgtggaagtg gtgaaagacc agtttcatac tcatgagggg tacatgatgg atttgacagc
1381 ccatcagggc cgggttggta atattctaca attgggaagt aagctgattg gaacaggaaa
1441 attatcagaa gatgaagaaa ctgaagtaca agagcagatg aatctcctaa attcaagatg
1501 ggaatgcctc agggtagcta gcatggaaaa acaaagcaat ttacatagag ttttaatgga
1561 tctccagaat cagaaactga aagagttgaa tgactggcta acaaaaacag aagaaagaac
1621 aaggaaaatg gaggaagagc ctcttggacc tgatcttgaa gacctaaaac gccaagtaca
1681 acaacataag gtgcttcaag aagatctaga acaagaacaa gtcagggtca attctctcac
1741 tcacatggtg gtggtagttg atgaatctag tggagatcac gcaactgctg ctttggaaga
1801 acaacttaag gtattgggag atcgatgggc aaacatctgt agatggacag aagaccgctg
1861 ggttctttta caagacatcc ttctcaaatg gcaacgtctt actgaagaac agtgcctttt
1921 tagtgcatgg ctttcagaaa aagaagatgc agtgaacaag attcacacaa ctggctttaa
1981 agatcaaaat gaaatgttat caagtcttca aaaactggcc gttttaaaag cggatctaga
2041 aaagaaaaag caatccatgg gcaaactgta ttcactcaaa caagatcttc tttcaacact
2101 gaagaataag tcagtgaccc agaagacgga agcatggctg gataactttg cccggtgttg
2161 ggataattta gtccaaaaac ttgaaaagag tacagcacag atttcacagg ctgtcaccac
2221 cactcagcca tcactaaaac agacaactgt aatggaaaca gtaactacgg tgaccacaag
2281 ggaacagatc ctggtaaagc atgctcaaga ggaacttcca ccaccacctc cccaaaagaa
2341 gaggcagatt actgtggatt ctgaaattag gaaaaggttg gatgttgata taactgaact
2401 tcacagctgg attactcgct cagaagctgt gttgcagagt cctgaatttg caatctttcg
2461 gaaggaaggc aacttctcag acttaaaaga aaaagtcaat gccatagagc gagaaaaagc
2521 tgagaagttc agaaaactgc aagatgccag cagatcagct caggccctgg tggaacagat
2581 ggtgaatgag ggtgttaatg cagatagcat caaacaagcc tcagaacaac tgaacagccg
2641 gtggatcgaa ttctgccagt tgctaagtga gagacttaac tggctggagt atcagaacaa
2701 catcatcgct ttctataatc agctacaaca attggagcag atgacaacta ctgctgaaaa
2761 ctggttgaaa atccaaccca ccaccccatc agagccaaca gcaattaaaa gtcagttaaa
2821 aatttgtaag gatgaagtca accggctatc aggtcttcaa cctcaaattg aacgattaaa
2881 aattcaaagc atagccctga aagagaaagg acaaggaccc atgttcctgg atgcagactt
2941 tgtggccttt acaaatcatt ttaagcaagt cttttctgat gtgcaggcca gagagaaaga
3001 gctacagaca atttttgaca ctttgccacc aatgcgctat caggagacca tgagtgccat
3061 caggacatgg gtccagcagt cagaaaccaa actctccata cctcaactta gtgtcaccga
3121 ctatgaaatc atggagcaga gactcgggga attgcaggct ttacaaagtt ctctgcaaga
3181 gcaacaaagt ggcctatact atctcagcac cactgtgaaa gagatgtcga agaaagcgcc
3241 ctctgaaatt agccggaaat atcaatcaga atttgaagaa attgagggac gctggaagaa
3301 gctctcctcc cagctggttg agcattgtca aaagctagag gagcaaatga ataaactccg
3361 aaaaattcag aatcacatac aaaccctgaa gaaatggatg gctgaagttg atgtttttct
3421 gaaggaggaa tggcctgccc ttggggattc agaaattcta aaaaagcagc tgaaacagtg
3481 cagactttta gtcagtgata ttcagacaat tcagcccagt ctaaacagtg tcaatgaagg
3541 tgggcagaag ataaagaatg aagcagagcc agagtttgct tcgagacttg agacagaact
```

FIGURE 23 (cont.)

```
3601 caaagaactt aacactcagt gggatcacat gtgccaacag gtctatgcca gaaaggaggc
3661 cttgaaggga ggtttggaga aaactgtaag cctccagaaa gatctatcag agatgcacga
3721 atggatgaca caagctgaag aagagtatct tgagagagat tttgaatata aaactccaga
3781 tgaattacag aaagcagttg aagagatgaa gagagctaaa gaagaggccc aacaaaaaga
3841 agcgaaagtg aaactcctta ctgagtctgt aaatagtgtc atagctcaag ctccacctgt
3901 agcacaagag gccttaaaaa aggaacttga aactctaacc accaactacc agtggctctg
3961 cactaggctg aatgggaaat gcaagacttt ggaagaagtt tgggcatgtt ggcatgagtt
4021 attgtcatac ttggagaaag caaacaagtg gctaaatgaa gtagaattta aacttaaaac
4081 cactgaaaac attcctggcg gagctgagga aatctctgag gtgctagatt cacttgaaaa
4141 tttgatgcga cattcagagg ataacccaaa tcagattcgc atattggcac agaccctaac
4201 agatggcgga gtcatggatg agctaatcaa tgaggaactt gagacattta attctcgttg
4261 gagggaacta catgaagagg ctgtaaggag gcaaaagttg cttgaacaga gcatccagtc
4321 tgcccaggag actgaaaaat ccttacactt aatccaggag tccctcacat tcattgacaa
4381 gcagttggca gcttatattg cagacaaggt ggacgcagct caaatgcctc aggaagccca
4441 gaaaatccaa tctgatttga caagtcatga gatcagttta gaagaaatga agaaacataa
4501 tcaggggaag gaggctgccc aaagagtcct gtctcagatt gatgttgcac agaaaaaatt
4561 acaagatgtc tccatgaagt ttcgattatt ccagaaacca gccaattttg agctgcgtct
4621 acaagaaagt aagatgattt tagatgaagt gaagatgcac ttgcctgcat tggaaacaaa
4681 gagtgtggaa caggaagtag tacagtcaca gctaaatcat tgtgtgaact tgtataaaag
4741 tctgagtgaa gtgaagtctg aagtggaaat ggtgataaag actggacgtc agattgtaca
4801 gaaaaagcag acggaaaatc ccaaagaact tgatgaaaga gtaacagctt tgaaattgca
4861 ttataatgag ctgggagcaa aggtaacaga aagaaagcaa cagttggaga aatgcttgaa
4921 attgtcccgt aagatgcgaa aggaaatgaa tgtcttgaca gaatggctgg cagctacaga
4981 tatggaattg acaaagagat cagcagttga aggaatgcct agtaatttgg attctgaagt
5041 tgcctgggga aaggctactc aaaaagagat tgagaaacag aaggtgcacc tgaagagtat
5101 cacagaggta ggagaggcct tgaaaacagt tttgggcaag aaggagacgt tggtggaaga
5161 taaactcagt cttctgaata gtaactggat agctgtcacc tcccgagcag aagagtggtt
5221 aaatcttttg ttggaatacc agaaacacat ggaaactttt gaccagaatg tggaccacat
5281 cacaaagtgg atcattcagg ctgacacact tttggatgaa tcagagaaaa agaaacccca
5341 gcaaaaagaa gacgtgctta agcgtttaaa ggcagaactg aatgacatac gcccaaaggt
5401 ggactctaca cgtgaccaag cagcaaactt gatggcaaac cgcggtgacc actgcaggaa
5461 attagtagag ccccaaatct cagagctcaa ccatcgattt gcagccattt cacacagaat
5521 taagactgga aaggcctcca ttcctttgaa ggaattggag cagtttaact cagatataca
5581 aaaattgctt gaaccactgg aggctgaaat tcagcagggg gtgaatctga aagaggaaga
5641 cttcaataaa gatatgaatg aagacaatga gggtactgta aaagaattgt tgcaaagagg
5701 agacaactta caacaaagaa tcacagatga gagaaagaga gaggaaataa agataaaaca
5761 gcagctgtta cagacaaaac ataatgctct caaggatttg aggtctcaaa gaagaaaaaa
5821 ggctctagaa atttctcatc agtggtatca gtacaagagg caggctgatg atctcctgaa
5881 atgcttggat gacattgaaa aaaaattagc cagcctacct gagcccagag atgaaaggaa
5941 aataaaggaa attgatcggg aattgcagaa gaagaaagag gagctgaatg cagtgcgtag
6001 gcaagctgag ggcttgtctg aggatggggc cgcaatggca gtggagccaa ctcagatcca
6061 gctcagcaag cgctggcggg aaattgagag caaatttgct cagtttcgaa gactcaactt
6121 tgcacaaatt cacactgtcc gtgaagaaac gatgatggtg atgactgaag acatgccttt
6181 ggaaatttct tatgtgcctt ctacttattt gactgaaatc actcatgtct cacaagccct
6241 attagaagtg gaacaacttc tcaatgctcc tgacctctgt gctaaggact ttgaagatct
6301 ctttaagcaa gaggagtctc tgaagaatat aaaagatagt ctacaacaaa gctcaggtcg
6361 gattgacatt attcatagca agaagacagc agcattgcaa agtgcaacgc ctgtggaaag
6421 ggtgaagcta caggaagctc tctcccagct tgatttccaa tgggaaaaag ttaacaaaat
6481 gtacaaggac cgacaagggc gatttgacag atctgttgag aaatggcggc gttttcatta
6541 tgatataaag atatttaatc agtggctaac agaagctgaa cagtttctca gaaagacaca
6601 aattcctgag aattgggaac atgctaaata caaatggtat cttaaggaac tccaggatgg
6661 cattgggcag cggcaaactg ttgtcagaac attgaatgca actggggaag aaataattca
6721 gcaatcctca aaaacagatg ccagtattct acaggaaaaa ttgggaagcc tgaatctgcg
6781 gtggcaggag gtctgcaaac agctgtcaga cagaaaaaag aggctagaag aacaaaagaa
6841 tatcttgtca gaatttcaaa gagatttaaa tgaatttgtt ttatggttgg aggaagcaga
6901 taacattgct agtatcccac ttgaacctgg aaaagagcag caactaaaag aaaagcttga
6961 gcaagtcaag ttactggtgg aagagttgcc cctgcgccag ggaattctca aacaattaaa
7021 tgaaactgga ggacccgtgc ttgtaagtgc tcccataagc ccagaagagc aagataaact
7081 tgaaaataag ctcaagcaga caaatctcca gtggataaag gtttccagag ctttacctga
7141 gaaacaagga gaaattgaag ctcaaataaa agaccttgag cagcttgaaa aaaagcttga
7201 agaccttgaa gagcagttaa atcatctgct gctgtggtta tctcctatta ggaatcagtt
7261 ggaaatttat aaccaaccaa accaagaagg accatttgac gttcaggaaa ctgaaatagc
7321 agttcaagct aaacaaccgg atgtggaaga gattttgtct aaagggcagc atttgtacaa
7381 ggaaaaacca gccactcagc cagtgaagag gaagttagaa gatctgagct ctgagtggaa
```

FIGURE 23 (cont.)

```
7441 ggcggtaaac cgtttacttc aagagctgag ggcaaagcag cctgacctag ctcctggact
7501 gaccactatt ggagcctctc ctactcagac tgttactctg gtgacacaac ctgtggttac
7561 taaggaaact gccatctcca aactagaaat gccatcttcc ttgatgttgg aggtacctgc
7621 tctggcagat ttcaaccggg cttggacaga acttaccgac tggctttctc tgcttgatca
7681 agttataaaa tcacagaggg tgatggtggg tgaccttgag gatatcaacg agatgatcat
7741 caagcagaag gcaacaatgc aggatttgga acagaggcgt ccccagttgg aagaactcat
7801 taccgctgcc caaaatttga aaaacaagac cagcaatcaa gaggctagaa caatcattac
7861 ggatcgaatt gaaagaattc agaatcagtg ggatgaagta caagaacacc ttcagaaccg
7921 gaggcaacag ttgaatgaaa tgttaaagga ttcaacacaa tggctggaag ctaaggaaga
7981 agctgagcag gtcttaggac aggccagagc caagcttgag tcatggaagg agggtcccta
8041 tacagtagat gcaatccaaa agaaaatcac agaaaccaag cagttggcca aagacctccg
8101 ccagtggcag acaaatgtag atgtggcaaa tgacttggcc ctgaaacttc tccgggatta
8161 ttctgcagat gataccagaa aagtccacat gataacagag aatatcaatg cctcttggag
8221 aagcattcat aaaagggtga gtgagcgaga ggctgctttg gaagaaactc atagattact
8281 gcaacagttc cccctggacc tggaaaagtt tcttgcctgg cttacagaag ctgaaacaac
8341 tgccaatgtc ctacaggatg ctacccgtaa ggaaaggctc ctagaagact ccaagggagt
8401 aaaagagctg atgaaacaat ggcaagacct ccaaggtgaa attgaagctc acacagatgt
8461 ttatcacaac ctggatgaaa acagccaaaa aatcctgaga tccctggaag gttccgatga
8521 tgcagtcctg ttacaaagac gtttggataa catgaacttc aagtggagtg aacttcggaa
8581 aaagtctctc aacattaggt cccatttgga agccagttct gaccagtgga agcgtctgca
8641 cctttctctg caggaacttc tggtgtggct acagctgaaa gatgatgaat taagccggca
8701 ggcacctatt ggaggcgact ttccagcagt tcagaagcag aacgatgtac atagggcctt
8761 caagagggaa ttgaaaacta aagaacctgt aatcatgagt actcttgaga ctgtacgaat
8821 atttctgaca gagcagcctt tggaaggact agagaaactc taccaggagc ccagagagct
8881 gcctcctgag gagagagccc agaatgtcac tcggcttcta cgaaagcagg ctgaggaggt
8941 caatactgag tgggaaaaat tgaacctgca ctccgctgac tggcagagaa aaatagatga
9001 gacccttgaa agactccagg aacttcaaga ggccacggat gagctggacc tcaagctgcg
9061 ccaagctgag gtgatcaagg gatcctggca gcccgtgggc gatctcctca ttgactctct
9121 ccaagatcac ctcgagaaag tcaaggcact tcgaggagaa attgcgcctc tgaaagagaa
9181 cgtgagccac gtcaatgacc ttgctcgcca gcttaccact ttgggcattc agctctcacc
9241 gtataacctc agcactctgg aagacctgaa caccagatgg aagcttctgc aggtggccgt
9301 cgaggaccga gtcaggcagc tgcatgaagc ccacagggac tttggtccag catctcagca
9361 ctttctttcc acgtctgtcc agggtccctg ggagagagcc atctcgccaa acaaagtgcc
9421 ctactatatc aaccacgaga ctcaaacaac ttgctgggac catcccaaaa tgacagagct
9481 ctaccagtct ttagctgacc tgaataatgt cagattctca gcttatagga ctgccatgaa
9541 actccgaaga ctgcagaagg ccctttgctt ggatctcttg agcctgtcag ctgcatgtga
9601 tgccttggac cagcacaacc tcaagcaaaa tgaccagccc atggatatcc tgcagattat
9661 taattgtttg accactattt atgaccgcct ggagcaagag cacaacaatt tggtcaacgt
9721 ccctctctgc gtggatatgt gtctgaactg gctgctgaat gtttatgata cgggacgaac
9781 agggaggatc cgtgtcctgt cttttaaaac tggcatcatt tccctgtgta aagcacattt
9841 ggaagacaag tacagatacc ttttcaagca agtggcaagt tcaacaggat tttgtgacca
9901 gcgcaggctg ggcctccttc tgcatgattc tatccaaatt ccaagacagt tgggtgaagt
9961 tgcatccttt gggggcagta acattgagcc aagtgtccgg agctgcttcc aatttgctaa
10021 taataagcca gagatcgaag cggccctctt cctagactgg atgagactgg aaccccagtc
10081 catggtgtgg ctgcccgtcc tgcacagagt ggctgctgca gaaactgcca agcatcaggc
10141 caaatgtaac atctgcaaag agtgtccaat cattggattc aggtacagga gtctaaagca
10201 ctttaattat gacatctgcc aaagctgctt tttttctggt cgagttgcaa aaggccataa
10261 aatgcactat cccatggtgg aatattgcac tccgactaca tcaggagaag atgttcgaga
10321 ctttgccaag gtactaaaaa acaaatttcg aaccaaaagg tattttgcga agcatccccg
10381 aatgggctac ctgccagtgc agactgtctt agagggggac aacatggaaa c
10471 gcctgcctcg tcccctcagc tttcacacga
10501 tgatactcat tcacgcattg aacattatgc tagcaggcta gcagaaatgg aaaacagcaa
10561 tggatcttat ctaaatgata gcatctctcc taatgagagc atagatgatg aacatttgtt
10621 aatccagcat tactgccaaa gtttgaacca ggactccccc ctgagccagc ctcgtagtcc
10681 tgcccagatc ttgatttcct tagagagtga ggaaagaggg gagctagaga gaatcctagc
10741 agatcttgag gaagaaaaca ggaatctgca agcagaatat gaccgtctaa agcagcagca
10801 cgaacataaa ggcctgtccc cactgccgtc ccctcctgaa atgatgccca cctctcccca
10861 gagtccccgg gatgctgagc tcattgctga ggccaagcta ctgcgtcaac acaaaggccg
10921 cctggaagcc aggatgcaaa tcctggaaga ccacaataaa cagctggagt cacagttaca
10981 caggctaagg cagctgctgg agcaacccca ggcagaggcc aaagtgaatg gcacaacggt
11041 gtcctctcct tctacctctc tacagaggtc cgacagcagt cagcctatgc tgctccgagt
11101 ggttggcagt caaacttcgg actccatggg tgaggaagat cttctcagtc ctccccagga
11161 cacaagcaca gggttagagg aggtgatgga gcaactcaac aactccttcc ctagttcaag
11221 aggaagaaat acccctggaa agccaatgag agaggacaca atgtaggaag tcttttccac
```

FIGURE 23 (cont.)

```
11281 atggcagatg atttgggcag agcgatggag tccttagtat cagtcatgac agatgaagaa
11341 ggagcagaat aaatgtttta caactcctga ttcccgcatg gtttttataa tattcataca
11401 acaaagagga ttagacagta agagtttaca agaaataaat ctatattttt gtgaagggta
11461 gtggtattat actgtagatt tcagtagttt ctaagtctgt tattgttttg ttaacaatgg
11521 caggttttac acgtctatgc aattgtacaa aaaagttata agaaaactac atgtaaaatc
11581 ttgatagcta aataacttgc catttcttta tatggaacgc attttgggtt gtttaaaaat
11641 ttataacagt tataaagaaa gattgtaaac taaagtgtgc tttataaaaa aaagttgttt
11701 ataaaaaccc ctaaaaacaa aacaaacaca cacacacaca catacacaca cacacacaaa
11761 actttgaggc agcgcattgt tttgcatcct tttggcgtga tatccatatg aaattcatgg
11821 ctttttcttt ttttgcatat taaagataag acttcctcta ccaccacacc aaatgactac
11881 tacacactgc tcatttgaga actgtcagct gagtggggca ggcttgagtt ttcatttcat
11941 atatctatat gtctataagt atataaatac tatagttata tagataaaga gatacgaatt
12001 tctatagact gacttttcc attttttaaa tgttcatgtc acatcctaat agaaagaaat
12061 tacttctagt cagtcatcca ggcttacctg cttggt
```

FIGURE 24

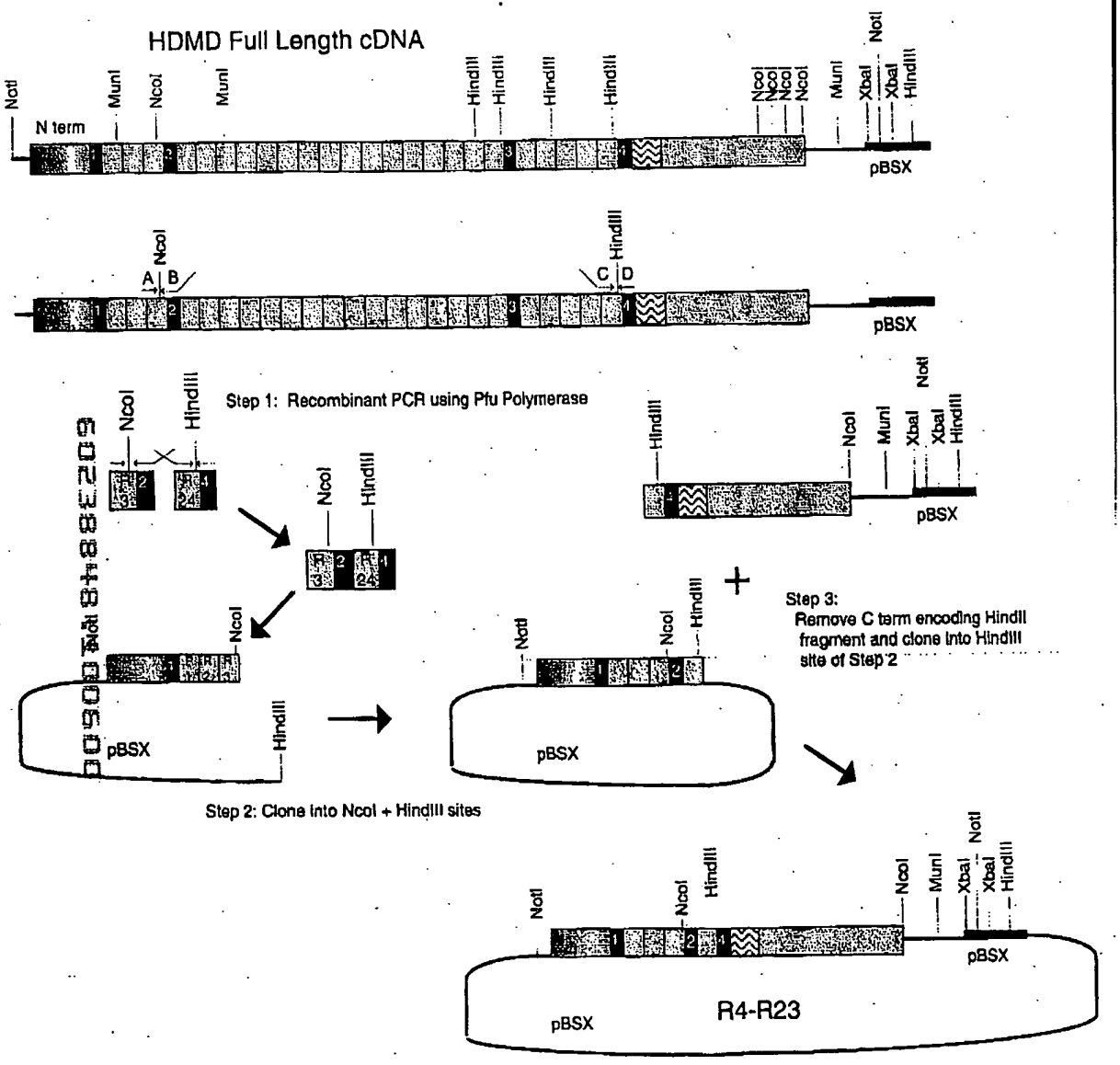

HDMD Full Length cDNA
NotI
MunI
NcoI
MunI
HindIII
HindIII
HindIII
HindIII
NcoI
NcoI
NcoI
MunI
XbaI
XbaI
HindIII
NotI
N term
pBSX
NcoI
HindIII
A
B
C
D
pBSX
Step 1: Recombinant PCR using Pfu Polymerase
NcoI
HindIII
HindIII
NcoI
MunI
XbaI
XbaI
HindIII
NotI
pBSX
NcoI
HindIII
NcoI
HindIII
pBSX
Step 3:
Remove C term encoding HindII
fragment and clone into HindIII
site of Step 2
NotI
NcoI
HindIII
pBSX
Step 2: Clone into NcoI + HindIII sites
NotI
NcoI
HindIII
NcoI
MunI
XbaI
XbaI
HindIII
NotI
pBSX
pBSX
R4-R23

FIGURE 25

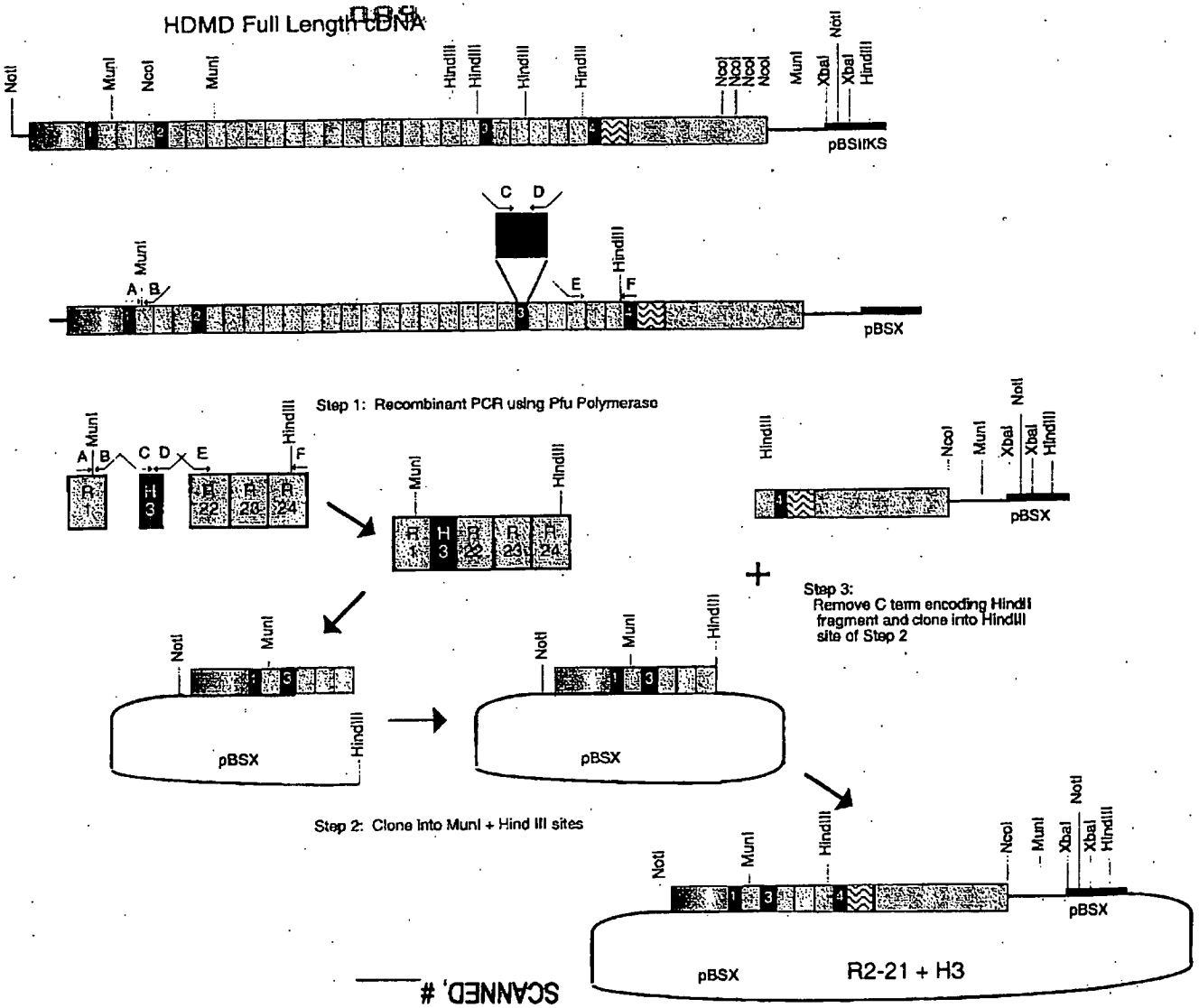


FIGURE 26

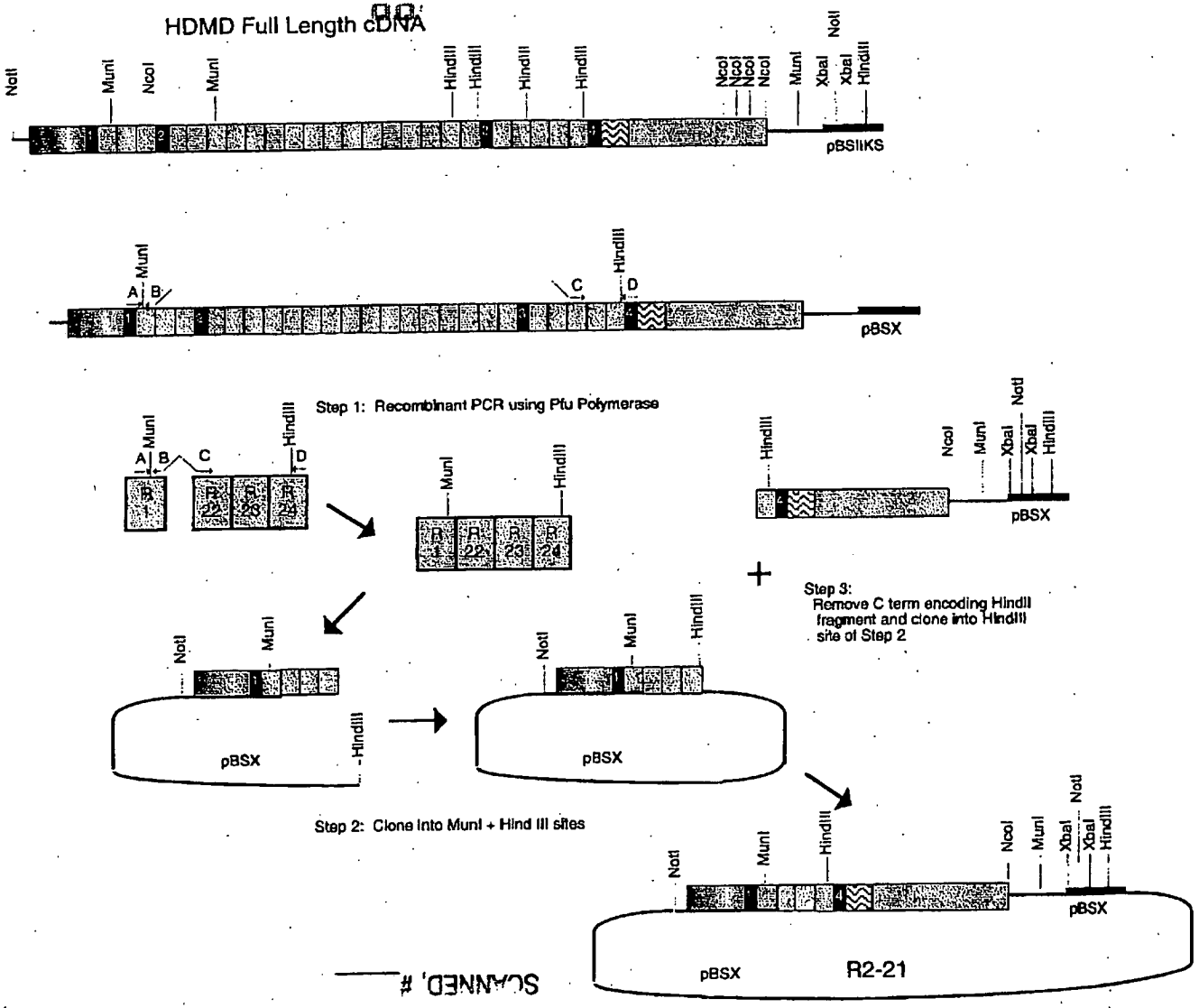


FIGURE 27

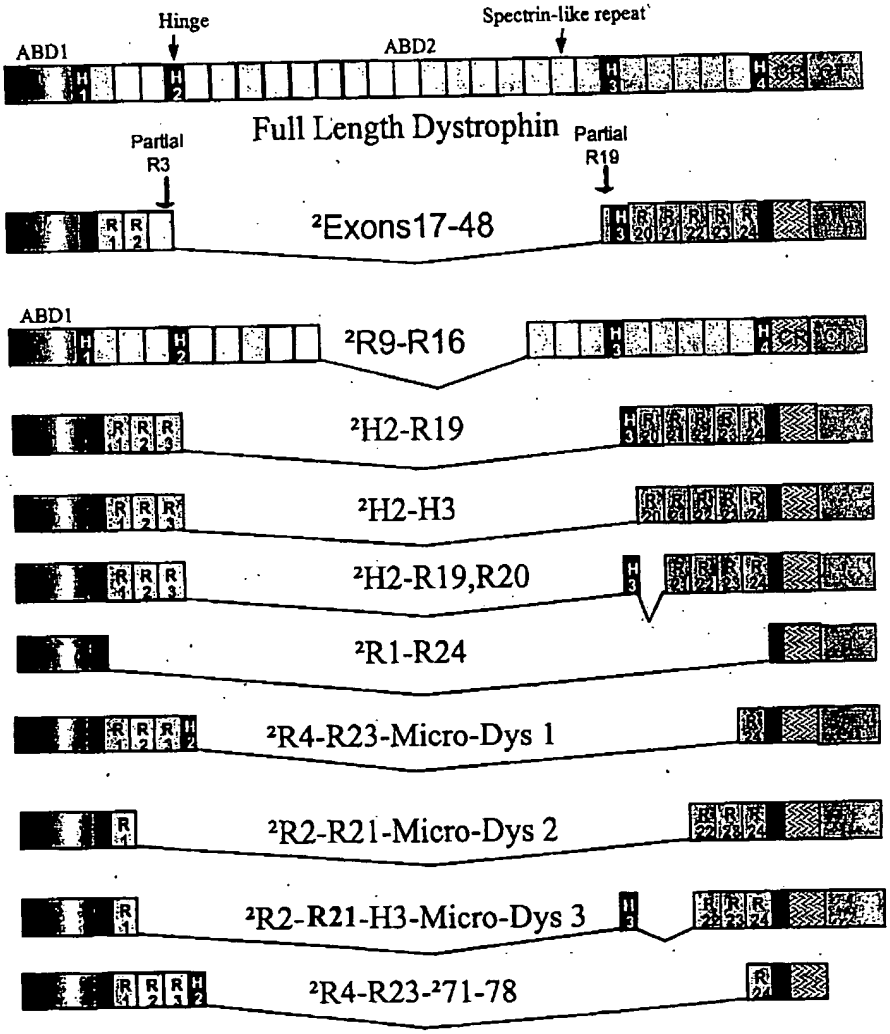


FIGURE 28

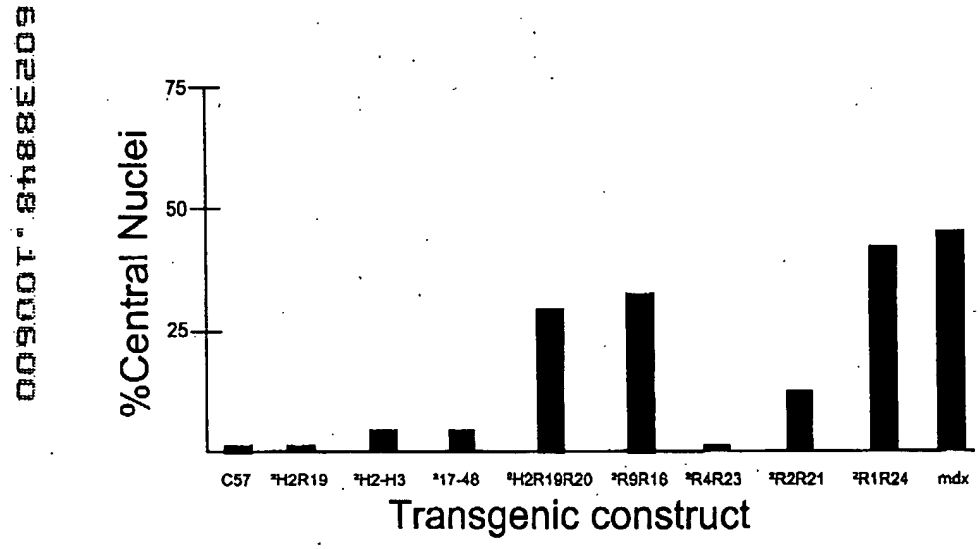

%Central Nuclei
75
50
25
C57
H2R19
H2-H3
17-48
H2R19R20
R9R16
R4R23
R2R21
R1R24
mdx
Transgenic construct

FIGURE 29

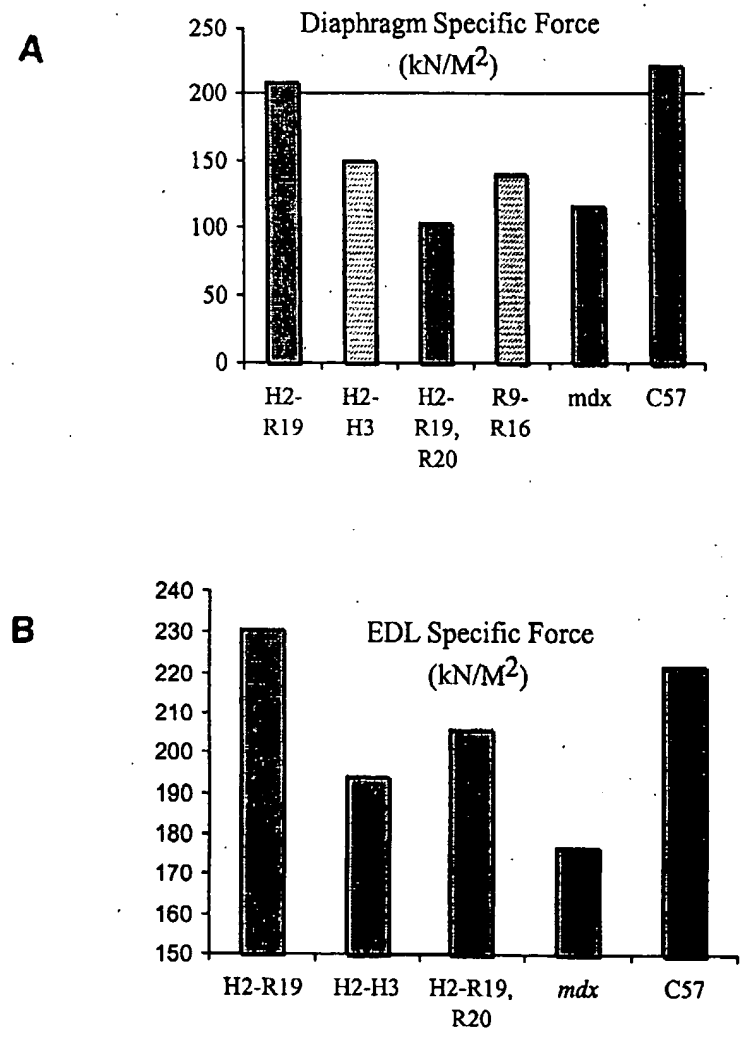
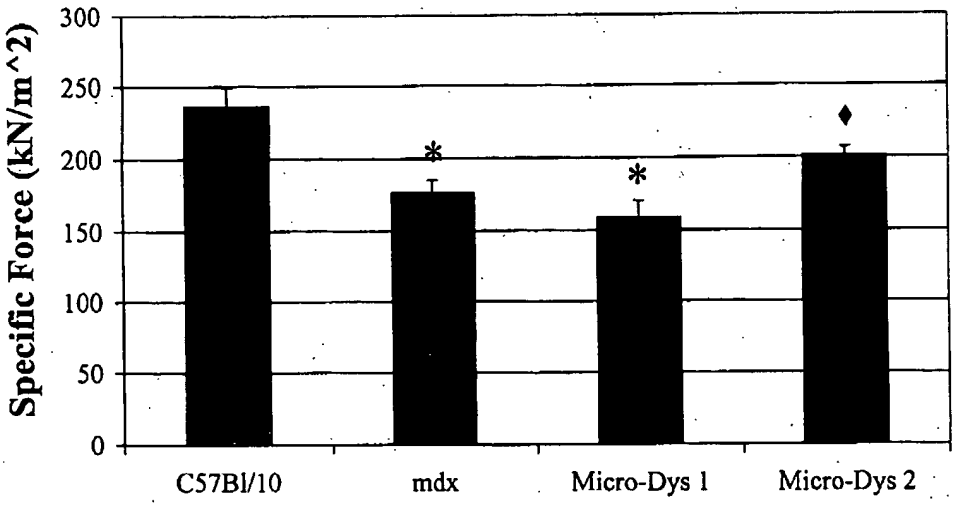


FIGURE 30

A. Specific Force in the TA Muscle



B. Specific Force in the Diaphragm

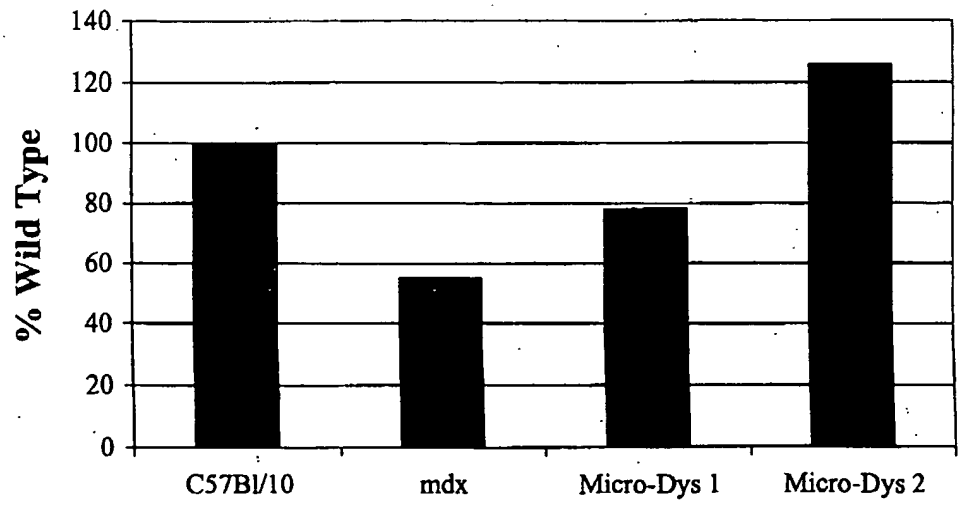


FIGURE 31

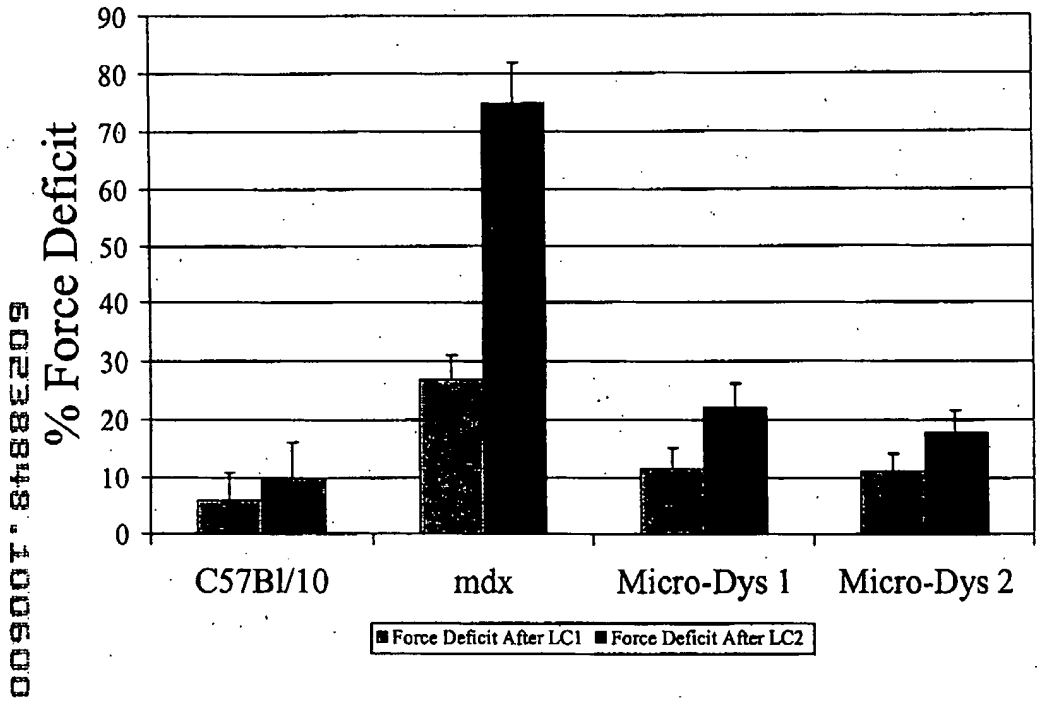

% Force Deficit
90
80
70
60
50
40
30
20
10
0
C57Bl/10
mdx
Micro-Dys 1
Micro-Dys 2
Force Deficit After LC1
Force Deficit After LC2

FIGURE 32

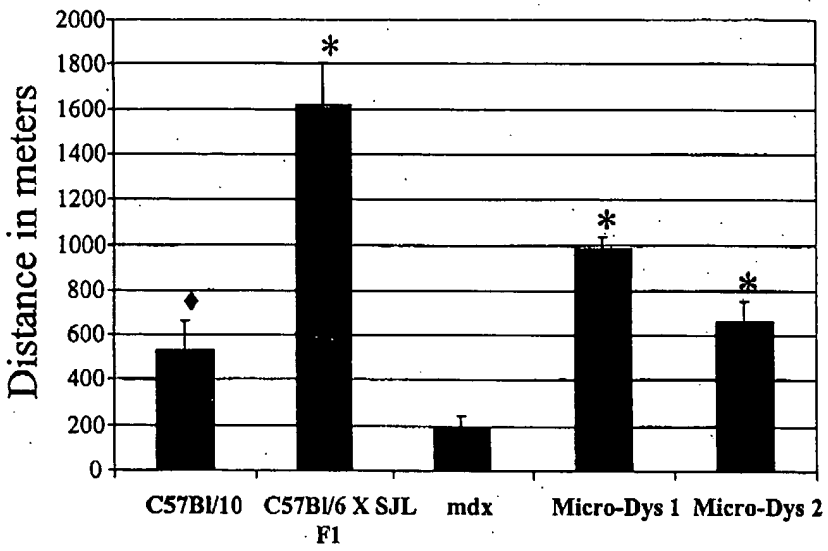


FIGURE 33

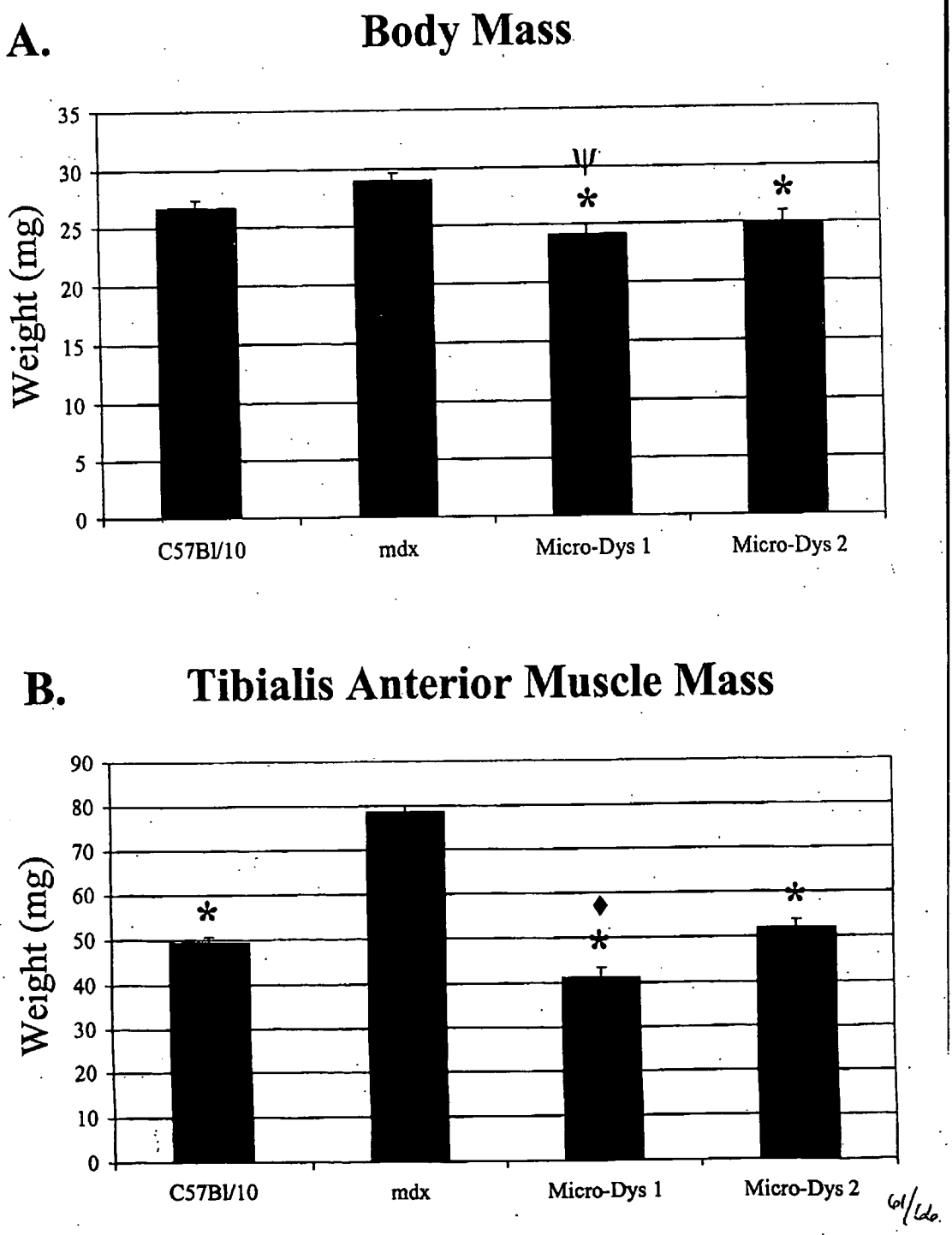

A.
Body Mass
Weight (mg)
35
30
25
20
15
10
5
0
C57Bl/10
mdx
Micro-Dys 1
Micro-Dys 2
B.
Tibialis Anterior Muscle Mass
Weight (mg)
90
80
70
60
50
40
30
20
10
0
C57Bl/10
mdx
Micro-Dys 1
Micro-Dys 2

FIGURE 34

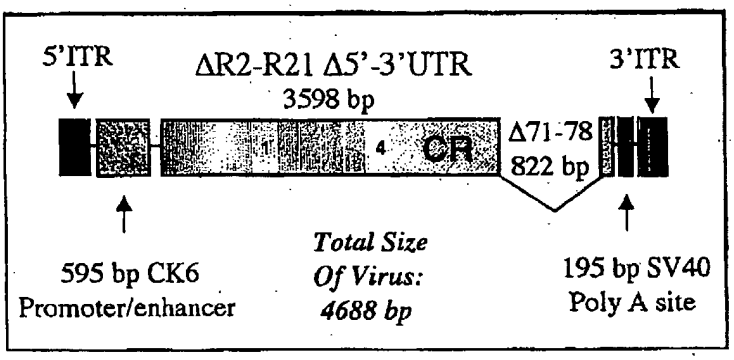


FIGURE 35

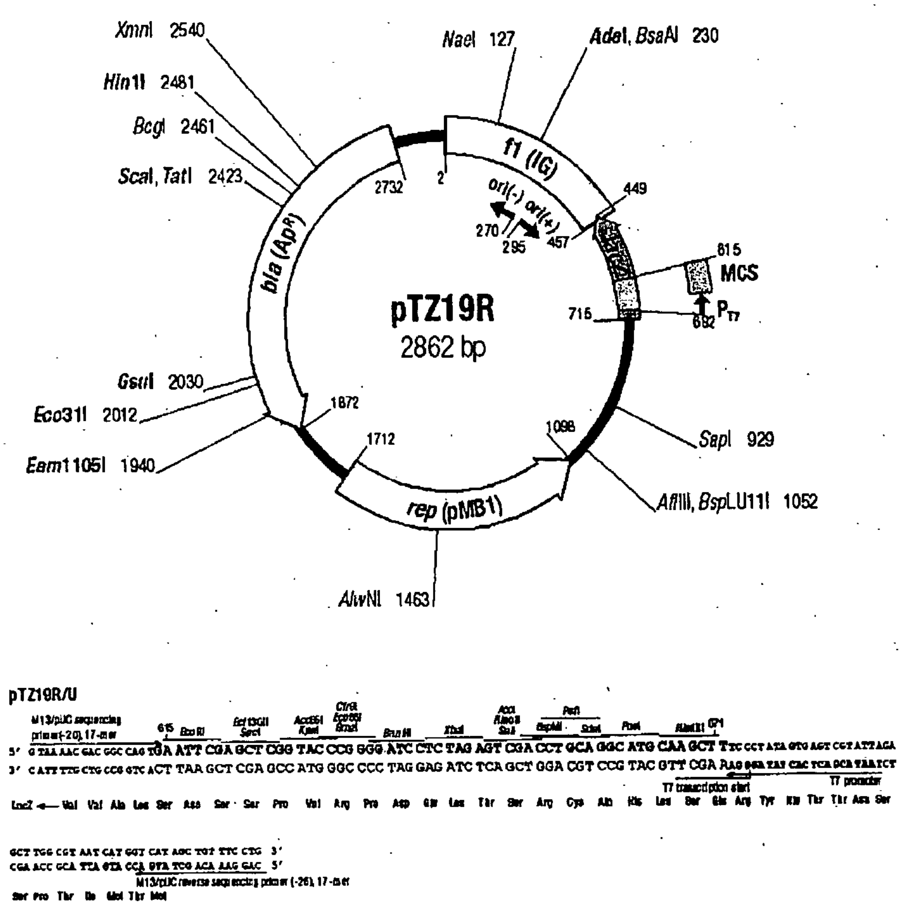

pTZ19R
2862 bp
XmnI 2540
HinII 2481
BcgI 2461
ScaI, TatI 2423
NaeI 127
AdeI, BsaAI 230
f1 (IG)
ori(-) ori(+)
270
295
457
449
2732
2
615
MCS
P_{T7}
692
715
bla (Ap^R)
GsuI 2030
Eco31I 2012
Eam1105I 1940
1872
1712
1098
SapI 929
AflIII, BspLU11I 1052
rep (pMB1)
AlwNI 1463
pTZ19R/U
M13/pUC sequencing primer (-20), 17-mer
T7 transcription start
T7 promoter
M13/pUC reverse sequencing primer (-26), 17-mer

FIGURE 36

SEQ ID NO: 87 wild type mouse enhancer) - CCACTA

2150 2200
CGGGTCTAGGCTGCCCATGTAAGGAGGCAAGGCCTGGGGACACCCGAGATGCCTGGTTATAATTAACCCAGACATGTGGCTGCCCCCCCCCCCCCCAACAC

2250 2300
CTGCTGCCTGAGCCTCACCCCCACCCCGGTGCCTGGGTCTTAGGCTCTGTACACCATGGAGGAGAAGCTCGCTCTAAAAATAACCCTGTCCCTGGTGGAT

SEQ ID NO: 88 ('2R' mouse mutant enhancer) - CCACTA

2150 2200
CGGGTCTAGGCTGCCCATGTAAGGAGGCAAGGCCTGGGGACACCCGAGATGCCTGGTTATAATTAACCCCAACACCTGCTGCCCCCCCCCCCCCCAACAC

2250 2300
CTGCTGCCTGAGCCTCACCCCCACCCCGGTGCCTGGGTCTTAGGCTCTGTACACCATGGAGGAGAAGCTCGCTCTAAAAATAACCCTGTCCCTGGTGGAT

SEQ ID NO: 89 ('S5' mouse mutant enhancer) - CCACTA

2150 2200
CGGGTCTAGGCTGCCCATGTAAGGAGGCAAGGCCTGGGGACACCCGAGATGCCTGGTTATAATTAACCCAGACATGTGGCTGCCCCCCCCCCCCCCAACAC

2250 2300
CTGCTGCCTGAGCCTGAGCGGTTACCCCACCCCGGTGCCTGGGTCTTAGGCTCTGTACACCATGGAGGAGAAGCTCGCTCTAAAAATAACCCTGTCCCTG

GTGGAT

SEQ ID NO: 90 ('2RS5' mouse mutant enhancer) - CCACTA

2150 2200
CGGGTCTAGGCTGCCCATGTAAGGAGGCAAGGCCTGGGGACACCCGAGATGCCTGGTTATAATTAACCCCAACACCTGCTGCCCCCCCCCCCCCCAACAC

2250 2300
CTGCTGCCTGAGCCTGAGCGGTTACCCCACCCCGGTGCCTGGGTCTTAGGCTCTGTACACCATGGAGGAGAAGCTCGCTCTAAAAATAACCCTGTCCCTG

GTGGAT

SEQ ID NO: 91 ('truncated 2RS5' mouse mutant enhancer) - CCACTA

2150 2200
CGGGTCTAGGCTGCCCATGTAAGGAGGCAAGGCCTGGGGACACCCGAGATGCCTGGTTATAATTAACCCCAACACCTGCTGCCCCCCCCCCCCCCAACAC

2250
CTGCTGCCTGAGCCTGAGCGGTTACCCCACCCCGGTGCCTGGGTCTTAGGCTCTGTACACCATGG

FIGURE 37

SEQ ID NO 92 (mouse promoter sequence, -944 to +7)

GTGGAGCAGCCTGCACTGGGCTTCTGGGAGAAACCAAACCGGGTTCTAACCTTTCAGCTACAGTTATTGCCTTTCCTGTAGATGGGCGACTACAGCCCCACC
CCCACCCCCGTCTCCTGTATCCTTCCTGGGCCTGGGGATCCTAGGCTTTCACTGGAAATTTCCCCCCAGGTGCTGTAGGCTAGAGTCACGGCTCCCAAGAAC
AGTGCTTGCCTGGCATGCATGGTTCTGAACCTCCAACTGCAAAAAATGACACATACCTTGACCCTTGGAAGGCTGAGGCAGGGGGATTGCCATGAGTGCAAA
GCCAGACTGGGTGGCATAGTTAGACCCTGTCTCAAAAAACCAAAAACAATTAAATAACTAAAGTCAGGCAAGTAATCCTACTCGGGAGACTGAGGCAGAGGG
ATTGTTACATGTCTGAGGCCAGCCTGGACTACATAGGGTTTCAGGCTAGCCCTGTCTACAGAGTAAGGCCCTATTTCAAAAACACAAACAAAATGGTTCTCC
CAGCTGCTAATGCTCACCAGGCAATGAAGCCTGGTGAGCATTAGCAATGAAGGCAATGAAGGAGGGTGCTGGCTACAATCAAGGCTGTGGGGGACTGAGGGC
AGGCTGTAACAGGCTTGGGGGCCAGGGCTTATACGTGCCTGGGACTCCCAAAGTATTACTGTTCCATGTTCCCGGCGAAGGGCCAGCTGTCCCCCGCCAGCT
AGACTCAGCACTTAGTTTAGGAACCAGTGAGCAAGTCAGCCCTTGGGGCAGCCCATACAAGGCCATGGGGCTGGGCAAGCTGCACGCCTGGGTCCGGGGTGG
GCACGGTGCCCGGGCAACGAGCTGAAAGCTCATCTGCTCTCAGGGGCCCCTCCCTGGGGACAGCCCCTCCTGGCTAGTCACACCCTGTAGGCTCCTCTATAT
AACCCAGGGGCACAGGGGCTGCCCCCGGGTCAC

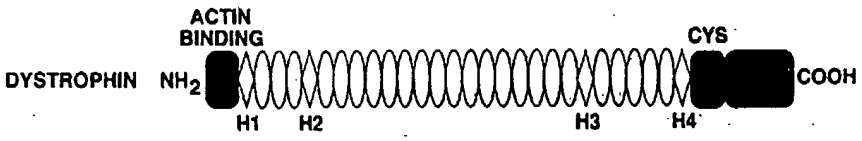
SEQ ID NO: 93 (mouse promoter sequence, -358 to +7)

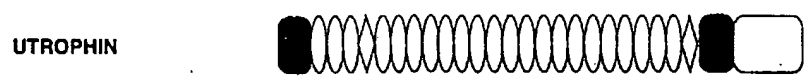
AATCAAGGCTGTGGGGGACTGAGGGCAGGCTGTAACAGGCTTGGGGGCCAGGGCTTATACGTGCCTGGGACTCCCAAAGTATTACTGTTCCATGTTCCCGGC
GAAGGGCCAGCTGTCCCCCGCCAGCTAGACTCAGCACTTAGTTTAGGAACCAGTGAGCAAGTCAGCCCTTGGGGCAGCCCATACAAGGCCATGGGGCTGGGC
AAGCTGCACGCCTGGGTCCGGGGTGGGCACGGTGCCCGGGCAACGAGCTGAAAGCTCATCTGCTCTCAGGGGCCCCTCCCTGGGGACAGCCCCTCCTGGCTA
GTCACACCCTGTAGGCTCCTCTATATAACCCAGGGGCACAGGGGCTGCCCCCGGGTCAC

SEQ ID NO:94 (mouse promoter sequence, -80 to +7)

CCTCCCTGGGGACAGCCCCTCCTGGCTAGTCACACCCTGTAGGCTCCTCTATATAACCCAGGGGCACAGGGGCTGCCCCCGGGTCAC

Figure 38





10-10-00

A/Prov

Express Mail Label No.: EL 658 77 1 US

PATENT

Attorney Docket No.: UM-04723

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 C.F.R. 1.53(b)(2).

Docket Number	UM-04723	Type a plus sign (+) inside this box →	+

INVENTOR(s) / APPLICANT(s)			
Last Name	First Name	Middle Initial	Residence (City and Either State or Foreign Country)
Chamberlain	**Jeffrey**	**S.**	**Ann Arbor, MI**
Harper	**Scott**	**Q.**	**Ann Arbor, MI**

TITLE OF THE INVENTION (280 Characters Max.)

Truncated Dystrophin Genes

CORRESPONDENCE ADDRESS

MEDLEN & CARROLL, LLP
220 Montgomery Street, Suite 2200
San Francisco, California 94104
United States of America

ENCLOSED APPLICATION PARTS *(Check All That Apply)*

X	Specification	Number of Pages 76	__	Small Entity Statement
X	Drawing(s)	*Number of Sheets* 66	X	Other (Specify): Power of Attorney (unexecuted)
			X	Other (Specify): Assignment (unexecuted)

METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT

Charge Account No. **08-1290** in the amount of **$75.00. An originally executed duplicate of this transmittal is enclosed for this purpose.**	FILING FEE AMOUNT ($)	**$150.00**
The Commissioner is hereby authorized to charge any deficiency in the payment of the required fee(s) and/or credit any overpayment to Deposit Account No.: **08-1290. An originally executed duplicate of this transmittal is enclosed for this purpose.**		

This invention was made by an agency of the United States Government under a contract with an agency of the United States Government.

__ No.

X Yes, the name of the U.S. Government agency and the Government contract number are: NIH R01AR40864-10.

Respectfully submitted,

Date: October 6, 2000

Jason R. Bond
Reg. No.: 45,439
MEDLEN & CARROLL, LLP
220 Montgomery Street, Suite 2200
San Francisco, California 94104
(415) 705-8410

__ Additional inventors are being named on separately numbered sheets attached hereto.

Express Mail Label No.: E[illegible]777 381 US

PATENT
Attorney Docket No.: **UM-04723**

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of: Jeffrey S. Chamberlain and Scott Q. Harper

For: **Truncated Dystrophin Genes**

Box Provisional Patent Application
Assistant Commissioner for Patents
Washington, D.C. 20231

CERTIFICATION UNDER 37 C.F.R. § 1.10

I hereby certify that this correspondence and the documents referred to as attached therein are being deposited with the United States Postal Service on **October 6, 2000**, in an envelope as "EXPRESS MAIL POST OFFICE TO ADDRESSEE" service under 37 C.F.R. § 1.10, Mailing Label Number **EL 658 777 381 US** addressed to: **Box Provisional Patent Application,** Assistant Commissioner for Patents, Washington, D.C. 20231.

Mary Ellen Waite

Mary Ellen Waite

60238848 100500

TRANSMITTAL COVER SHEET FOR FILING PROVISIONAL APPLICATION (37 C.F.R. § 1.51(2)(i))

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 C.F.R. 1.53(b)(2).

1. The following comprises the information required by 37 C.F.R. § 1.51(a)(2)(i)(A):

2. The name(s) of the inventor(s) is/are (37 C.F.R. § 1.51(a)(2)(i)(B)):

   Jeffrey S. Chamberlain
   Scott Q. Harper

3. Address(es) of the inventor(s), as numbered above (37 C.F.R. § 1.51(a)(2)(i)(C)):

   1209 Henry St., Ann Arbor, MI 48104
   2357 Stone Rd., Ann Arbor, MI

4. The title of the invention is (37 C.F.R. § 1.51(a)(2)(i)(D)):

   Truncated Dystrophin Genes

5. The name, registration, and telephone number of the attorney (*if applicable*) is (37 C.F.R. § 1.51(a)(2)(i)(E)):

   Jason R. Bond
   Reg. No.: 44,439
   Tel.: (415) 705-8410

(complete the following, if applicable)

X An unexecuted Power of Attorney accompanies this cover sheet.

6. The docket number used to identify this application is (37 C.F.R. § 1.51(a)(2)(i)(F)):

Docket No.: **UM-04723**

7. The correspondence address for this application is (37 C.F.R. § 1.51(a)(2)(i)(G)):

MEDLEN & CARROLL, LLP
220 Montgomery Street, Suite 2200
San Francisco, California 94104

8. Statement as to whether invention was made by an agency of the U.S. Government or under contract with an agency of the U.S. Government. (37 C.F.R. § 1.51(a)(2)(i)(H)):

This invention was made by an agency of the United States Government, or under contract with an agency of the United States Government.

___ No.

_X_ Yes.

The name of the U.S. Government agency and the Government contract number are: NIH R01AR40864-10.

9. Identification of documents accompanying this cover sheet:

A. Documents required by 37 C.F.R. § 1.51(a)(2)(ii)-(iii):

Specification: No. of pages 76

Drawings: No. of sheets 66

B. Additional documents:

_X_ Claims: No. of claims 30

_X_ Power of Attorney (unexecuted)

___ Small Entity Statement

_X_ Assignment (unexecuted)

___ Other

10. Fee

The filing fee for this provisional application, as set in 37 C.F.R. § 1.16(k), is $150.00, for other than a small entity, and $75.00, for a small entity.

___ Applicant is a small entity.

11. Small Entity Statement

___ The verified statement(s) that this is a filing by a small entity under 37 C.F.R. §§ 1.9 and 1.27 is(are) attached.

Express Mail Label No.: EI 777 381 US

PATENT
Attorney Docket No.: **UM-04723**

12. Fee payment being made at this time

__ Charge Account No. **08-1290** in the amount of **$75.00. An originally executed duplicate of this transmittal is enclosed for this purpose.**

13. Method of Fee Payment:

__X__ Check in the amount of **$150.00**

____ Charge Account No. **08-1290**, in the amount of **$75.00. A duplicate of this Cover Sheet is attached.**

__X__ Please charge Account No. **08-1290** for any fee deficiency. **A duplicate of this Cover Sheet is attached.**

Date: October 6, 2000

Jason R. Bond
Reg. No.: 45,439

MEDLEN & CARROLL, LLP
220 Montgomery Street, Suite 2200
San Francisco, California 94104
(415) 705-8410

60238848.100500

REQUEST FOR ACCESS TO AN ABANDONED APPLICATION UNDER 37 CFR 1.14

In re Application of

Application Number: 60/238848 Filed: 10-6-00

Bring completed form to:
File Information Unit
Crystal Plaza Three, Room 1D03
2021 South Clark Place
Arlington, VA
Telephone: (703) 305-8733

RECEIVED
MAY 06 2004
File Information Unit

Paper No. #2

I hereby request access under 37 CFR 1.14 to the application file record of the above-identified ABANDONED application, which is identified in, or to which a benefit is claimed, in the following document (as shown in the attachment):

United States Patent Application Publication No. 2003/0216332, page _____, line _____

United States Patent Number _____, column _____, line _____, or

WIPO Pub. No. _____, page _____, line _____.

Related Information about Access to Pending Applications (37 CFR 1.14):

Direct access to pending applications is not available to the public but copies may be available and may be purchased from the Office of Public Records upon payment of the appropriate fee (37 CFR 1.19(b)), as follows:

For published applications that are still pending, a member of the public may obtain a copy of:
- the file contents;
- the pending application as originally filed; or
- any document in the file of the pending application.

For unpublished applications that are still pending:

(1) If the benefit of the pending application is claimed under 35 U.S.C. 119(e), 120, 121, or 365 in another application that has: (a) issued as a U.S. patent, or (b) published as a statutory invention registration, a U.S. patent application publication, or an international patent application publication in accordance with PCT Article 21(2), a member of the public may obtain a copy of:
- the file contents;
- the pending application as originally filed; or
- any document in the file of the pending application.

(2) If the application is incorporated by reference or otherwise identified in a U.S. patent, a statutory invention registration, a U.S. patent application publication, or an international patent application publication in accordance with PCT Article 21(2), a member of the public may obtain a copy of:
- the pending application as originally filed.

Signature

Typed or printed name

Registration Number, if applicable

(703) 615-1679
Telephone Number

5-6-04
Date

RECEIVED
FOR PTO USE ONLY
MAY 06 2004
File Information Unit

Approved by: V. Boatman

Unit: _____

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.